IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

ASSISTANT COMMISSIONER FOR PATENTS
Washington, D.C. 20231

Atty's Docket Number: 6999.0005-01

Prior Application: PCT/GB98/00216

Art Unit: Not yet assigned

Examiner: Not yet assigned

SIR: This is a request for filing a

☒ Continuation ☐ Continuation-in-Part ☐ Divisional Application under 37 C.F.R. § 1.53(b) of pending prior application Serial No. PCT/GB98/00216 filed January 26, 1998 of Catherine Clare BLACKBURN, Ian Paul CHAMBERS, Alexander L. MEDVINSKY, Hitoshi NIWA, and Austin G. SMITH for DNA EXPRESSION IN TRANSFECTED CELLS AND ASSAYS CARRIED OUT IN TRANSFECTED CELLS

1. ☒ Enclosed is a complete copy of the prior application including the oath or Declaration and drawings, if any, as originally filed. I hereby verify that the attached papers are a true copy of prior application Serial No. PCT/GB98/00216 as originally filed on January 26, 1998.

2. ☐ Enclosed is a substitute specification under 37 C.F.R. § 1.125.

3. ☐ Cancel Claims _____.

4. ☒ A Preliminary Amendment is enclosed.

5. ☒ The filing fee is calculated on the basis of the claims existing in the prior application as amended at 3 and 4 above.

LAW OFFICES
FINNEGAN, HENDERSON,
FARABOW, GARRETT,
& DUNNER, L.L.P.
1300 I STREET, N. W.
WASHINGTON, D.C. 20005
202-408-4000

For	Number Filed	Number Extra	Rate	Basic Fee $760.00
Total Claims	50 -20=	30	x$ 18.00=	$ 540.00
Independent Claims	4 -3=	1	x$ 78.00=	78.00
Multiple Dependent Claim(s) (if applicable)			+$260.00=	
			Total =	1,458.00
		Reduction by ½ for filing by small entity		-
		TOTAL FILING FEE	=	1,458.00

6. ☒ A check in the amount of $ 1,458.00 to cover the filing fee is enclosed.

7. ☐ The Commissioner is hereby authorized to charge any fees which may be required including fees due under 37 C.F.R. § 1.16 and any other fees due under 37 C.F.R. § 1.17, or credit any overpayment during the pendency of this application to Deposit Account No. 06-0916.

8. ☒ Amend the specification by inserting before the first line, the sentence:

--This is a X continuation ☐ division of application Serial No. PCT/GB98/00216, filed January 24, 1997-- all of which are incorporated herein by reference.

9. ☐ New formal drawings are enclosed.

10. ☐ The prior application is assigned of record to:_____

_____.

11. ☐ Priority of application Serial No. 9701492.2, filed on January 24, 1997 in Great Britain (country) is claimed under 35 U.S.C. § 119. A certified copy

☐ is enclosed or ☐ is on file in the prior application.

12. ☐ A verified statement claiming small entity status

☐ is enclosed or ☐ is on file in the prior application.

LAW OFFICES
FINNEGAN, HENDERSON,
FARABOW, GARRETT,
& DUNNER, L.L.P.
1300 I STREET, N. W.
WASHINGTON, D. C. 20005
202-408-4000

13. ☒ The power of attorney in the prior application is to at least one of the following: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, L.L.P., Douglas B. Henderson, Reg. No. 20,291; Ford F. Farabow, Jr., Reg. No. 20,630; Arthur S. Garrett, Reg. No. 20,338; Donald R. Dunner, Reg. No. 19,073; Brian G. Brunsvold, Reg. No. 22,593; Tipton D. Jennings, IV, Reg. No. 20,645; Jerry D. Voight, Reg. No. 23,020; Laurence R. Hefter, Reg. No. 20,827; Kenneth E. Payne, Reg. No. 23,098; Herbert H. Mintz, Reg. No. 26,691; C. Larry O'Rourke, Reg. No. 26,014; Albert J. Santorelli, Reg. No. 22,610; Michael C. Elmer, Reg. No. 25,857; Richard H. Smith, Reg. No. 20,609; Stephen L. Peterson, Reg. No. 26,325; John M. Romary, Reg. No. 26,331; Bruce C. Zotter, Reg. No. 27,680; Dennis P. O'Reilley, Reg. No. 27,932; Allen M. Sokal, Reg. No. 26,695; Robert D. Bajefsky, Reg. No. 25,387; Richard L. Stroup, Reg. No. 28,478; David W. Hill, Reg. No. 28,220; Thomas L. Irving, Reg. No. 28,619; Charles E. Lipsey, Reg. No. 28,165; Thomas W. Winland, Reg. No. 27,605; Basil J. Lewris, Reg. No. 28,818; Martin I. Fuchs, Reg. No. 28,508; E. Robert Yoches, Reg. No. 30,120; Barry W. Graham, Reg. No. 29,924; Susan Haberman Griffen, Reg. No. 30,907; Richard B. Racine, Reg. No. 30,415; Thomas H. Jenkins, Reg. No. 30,857; Robert E. Converse, Jr., Reg. No. 27,432; Clair X. Mullen, Jr., Reg. No. 20,348; Christopher P. Foley, Reg. No. 31,354; John C. Paul, Reg. No. 30,413; David M. Kelly, Reg. No. 30,953; Kenneth J. Meyers, Reg. No. 25,146; Carol P. Einaudi, Reg. No. 32,220; Walter Y. Boyd, Jr., Reg. No. 31,738; Steven M. Anzalone, Reg. No. 32,095; Jean B. Fordis, Reg. No. 32,984; Roger D. Taylor, Reg. 28,992; Barbara C. McCurdy, Reg. No. 32,120; James K. Hammond, Reg. No. 31,964; Richard V. Burgujian, Reg. No. 31,744; J. Michael Jakes, Reg. No. 32,824; Thomas W. Banks, Reg. No. 32,719; Christopher P. Isaac, Reg. No. 32,616; Bryan C. Diner, Reg. No. 32,409; M. Paul Barker, Reg. No. 32,013; Andrew Chanho Sonu, Reg. No. 33,457; David S. Forman, Reg. No. 33,694; Vincent P. Kovalick, Reg. No. 32,867; James W. Edmondson, Reg. No. 33,871; Michael R. McGurk, Reg. No. 32,045; Joann M. Neth, Reg. No. 36,363; Gerson S. Panitch, Reg. No. 33,751; Cheri M. Taylor, Reg. No. 33,216; Charles E. Van Horn, Reg. No. 40,266; Linda A. Wadler, Reg. No. 33,218; Jeffrey A. Berkowitz, Reg. No. 36,743; Michael R. Kelly, Reg. No. 33, 921; and James B. Monroe, Reg. No. 33,971.

14. ☐ The power appears in the original declaration of the prior application.

LAW OFFICES
FINNEGAN, HENDERSON,
FARABOW, GARRETT,
& DUNNER, L.L.P.
1300 I STREET, N.W.
WASHINGTON, D.C. 20005
202-408-4000

15. ☐ Since the power does not appear in the original declaration, a copy of the power in the prior application is enclosed.

16. ☒ Please address all correspondence to FINNEGAN, HENDERSON, FARABOW, GARRETT and DUNNER, L.L.P., 1300 I Street, N.W., Washington, D.C. 20005-3315.

17. ☐ Recognize as associate attorney _____

_____

(name, address & Reg. No.)

18. ☐ Also enclosed is an Information Disclosure Statement _____

_____

PETITION FOR EXTENSION. If any extension of time is necessary for the filing of this application, including any extension in the parent application, serial no. _____, filed _____, for the purpose of maintaining copendency between the parent application and this application, and such extension has not otherwise been requested, such an extension is hereby requested, and the Commissioner is authorized to charge necessary fees for such an extension to our Deposit Account No. 06-0916. A duplicate copy of this paper is enclosed for use in charging the deposit account.

FINNEGAN, HENDERSON, FARABOW,
GARRETT & DUNNER, L.L.P.

By: Leslie McDonell
Leslie A. McDonell
Reg. No.: 34,872

Date: July 23, 1999

LAW OFFICES
FINNEGAN, HENDERSON,
FARABOW, GARRETT,
& DUNNER, L.L.P.
1300 I STREET, N. W.
WASHINGTON, D. C. 20005
202-408-4000

PATENT
Attorney Docket No. 6999.0005-01

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Catherine BLACKBURN et al.	)	
Continuation of PCT/GB98/00216	)	Group Art Unit: Not yet assigned
Filed: July 23, 1999 (Concurrently Herewith)	)	Examiner: Not yet assigned
For: **DNA EXPRESSION IN TRANSFECTED CELLS AND ASSAYS CARRIED OUT IN TRANSFECTED CELLS**	)	

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

PRELIMINARY AMENDMENT

Prior to calculation of the claims fees in the above-identified application, please enter the following amendments.

IN THE CLAIMS:

Claim 3, delete "1 or".

Claim 4, delete "any of Claims 1-3" and insert therefor -- Claim 1 --.

Claim 5, delete "any of Claims 1-4" and insert therefor -- Claim 1 --.

Claim 6, delete "any of Claims 1-5" and insert therefor -- Claim 1 --.

Claim 8, delete "any preceding claim" and insert therefor -- Claim 1 --.

Claim 9, delete "any preceding claim" and insert therefor -- Claim 1 --.

Claim 11, delete "any preceding claim" and insert therefor -- Claim 1 --.

Claim 12, delete "any preceding claim" and insert therefor -- Claim 1 --.

Claim 13, delete "any of Claims 1-11" and insert therefor -- Claim 1 --.

LAW OFFICES
FINNEGAN, HENDERSON,
FARABOW, GARRETT
& DUNNER, L.L.P.
1300 I STREET, N W
WASHINGTON, DC 20005
202-408-4000

Claim 14, delete "any preceding claims" and insert therefor -- Claim 1 --.

Claim 15, delete "any preceding claim" and insert therefor -- Claim 1 --.

Claim 16, delete "any preceding claim" and insert therefor -- Claim 1 --.

Claim 19, delete "17 or".

Claim 20, delete "any of Claims 17 to 19" and insert therefor -- Claim 17 --.

Claim 21, delete "any of Claims 17 to 20" and insert therefor -- Claim 17 --.

Claim 22, delete "any of Claims 17 to 21" and insert therefor -- Claim 17 --.

Claim 24, delete "any of Claims 17 to 23" and insert therefor -- Claim 17 --.

Claim 25, delete "any of Claims 17 to 23" and insert therefor -- Claim 17 --.

Claim 26, delete "any of Claims 17 to 25" and insert therefor -- Claim 17 --.

Claim 27, delete "any of Claims 17 to 26" and insert therefor -- Claim 17 --.

Cancel claim 28.

Claim 29, delete "any of Claims 17 to 27" and insert therefor -- Claim 17 --.

Claim 32, delete "any of Claims 29 to 31" and insert therefor -- Claim 29 --.

Claim 35, delete "or 34".

Claim 36, delete "any of Claims 33 to 35" and insert therefor -- Claim 33 --.

Claim 39, delete "or 38".

Claim 40, delete "any of Claims 37 to 39" and insert therefor -- Claim 37 --.

Claim 41, delete "any of Claims 37 to 40" and insert therefor -- Claim 37 --.

Claim 42, delete "any of Claims 37 to 41" and insert therefor -- Claim 37 --.

LAW OFFICES
FINNEGAN, HENDERSON,
FARABOW, GARRETT
& DUNNER, L.L.P.
1300 I STREET, N W
WASHINGTON, DC 20005
202-408-4000

Claim 44, delete "any of Claims 37 to 43" and insert therefor -- Claim 37 --.

Claim 45, delete "any of Claims 37 to 44" and insert therefor -- Claim 37 --.

Claim 47, delete "any of Claims 37 to 46" and insert therefor -- Claim 37 --.

Claim 49, delete "any of Claims 37 to 48" and insert therefor -- Claim 37 --.

Claim 50, delete "any of Claims 37 to 48" and insert therefor -- Claim 37 --.

REMARKS

Applicants have amended the claims to correct improper multiple dependencies. No new matter is added.

If any fees are required for entry of this Preliminary Amendment, please charge those fees to Deposit Account No. 06-0916.

Respectfully submitted,

FINNEGAN, HENDERSON, FARABOW,
GARRETT & DUNNER, L.L.P.

Dated: July 23, 1999

By: Leslie McDonell
Leslie A. McDonell
Reg. No. 34,872

LAW OFFICES
FINNEGAN, HENDERSON,
FARABOW, GARRETT
& DUNNER, L.L.P.
1300 I STREET, N W
WASHINGTON, DC 20005
202-408-4000

DNA Expression In Transfected Cells And Assays Carried Out In Transfected Cells

This invention relates to methods of expressing DNA in cells, to vectors for expression of DNA in cells and to transfected cells. The invention also relates to assays carried out in transfected cells or differentiated derivatives of such cells. In particular the invention relates to transfection of and expression of DNA in embryonic stem (ES) cells.

The wealth of sequence information now becoming available from the genome projects demands the development of new, high throughput systems for functional analysis. A powerful route to discovering and characterising genes involved in determination and differentiation in mammals is potentially available via the genetic manipulation of ES cells *in vitro*.

ES cells, which are derived from the pluripotential inner cell mass (ICM) of the preimplantation mouse embryo (2,3), retain the capacity for multilineage differentiation both *in vitro* (4,5) and *in vivo* (6,7). In principle, therefore, gene products which influence developmental decisions should be assayable in ES cell culture systems, whatever the source of the cells. However, there are major difficulties in analysing cDNA function by ES cell transfection. The frequency of isolating stable transfectants is low ($<10^{-4}$ by electroporation, calcium phosphate co-precipitation or lipofection) and the great majority of transfectants show heterogeneous and unstable expression.

These problems are particularly significant in the case of cDNAs whose expression causes differentiation because differentiated ES cell progeny do not generally proliferate. In such cases transfectants may still be isolated but transgene expression will be minimal.

Episomal vectors have been used for functional screening in other cell types in order to increase the frequency of stable transfection and to achieve reliable transgene expression. However, previously described episomal vectors, for example based on Epstein-Barr virus (EBV) or bovine papilloma virus (BPV), have limitations both in

host cell range and maintenance during long-term culture.

A modified extrachromosomal vector is known based on the replication system of murine polyoma virus (8). This plasmid, pMGD20neo, can be stably maintained as an episome in ES cells during long term culture. Importantly, the low levels of large T protein produced have no overt effect on the growth or differentiation properties of the ES cells (8,9). It is also known to maintain simultaneously with pMGD20neo a second episomal vector. Expression from the second vector was not possible hence pMGD20neo was used for cDNA expression. However, this vector already comprises two expression cassettes, one each for large T antigen and the neo selectable marker so its size constrains its use for expression of a third cassette containing a cDNA.

It is an object of the invention to provide a vector for transfection of and expression of DNA within a cell and a method of expressing DNA in a cell that overcomes or at least ameliorates the disadvantages identified in the art. An object of at least the preferred embodiments of the invention is to achieve, in a transfected cell, expression that is more stable and more homogenous than hitherto attainable. Further objects of preferred embodiments of the invention are to provide a method of expressing a DNA in an embryonic cell in a more stable and more homogenous manner than hitherto attainable, and to provide for stable transfection of embryonic cells at a higher frequency than can be obtained using conventional vectors.

The invention is based upon the maintenance of a vector within a cell, wherein maintenance of the vector is dependant upon the continued presence within the cell of a certain factor and wherein that factor is not expressed by the vector but is produced in or present in the cell in an amount sufficient to maintain the vector.

Accordingly the invention provides a transfection and expression method comprising, in a cell that expresses or will express a replication factor, introducing a vector dependant upon that replication factor. Thus, in a first aspect, the invention provides a method of expressing a DNA in a cell, comprising:

(a) (i) transfecting the cell with a first vector that expresses a

replication factor; or

(ii) otherwise obtaining a cell that expresses or will express the replication factor;

and

(b) transfecting the cell with a second vector, wherein

(i) the second vector contains a DNA, or is adapted to receive a DNA, in operative combination with a promoter for expression of the DNA; and

(ii) extrachromosomal replication of the second vector is dependant upon presence within the cell of the replication factor.

The replication factor is optionally non-toxic to the cell. Alternatively, the replication factor is toxic to the cell at high levels of expression but at low levels of expression is substantially non-toxic to the cell but at these low levels is present in sufficient amount to enable replication of the second vector.

Further, the replication factor preferably does not alter the ability of the cell to differentiate or proliferate, and may thus be regarded as being neutral to the cell phenotype. This enables the activities of the product of a cDNA to be investigated over a long time period and many cell generations without having to take account of possible interfering effects of the replication factor present within the cell. Again, the replication factor may be phenotype-neutral at all levels or may be neutral at a low level which is nevertheless a sufficient level to maintain the second vector within the cell.

The invention is of application to all cell types for which there exists, whether from a natural or synthetic source, a replication factor capable of maintaining in that cell type an episomal vector. The vector is preferably stably maintained, meaning it is maintained over a number of cell generations, and at least over 3 generations. The cell is preferably selected from the group consisting of mammalian cells, in particular primate cells or murine cells, and avian cells. It is further preferred that the cell is an embryonic cell, in particular an ES, EC (embryonic carcinoma) or EG (embryonic

gonadal) cell, or differentiated progeny of any such cell.

While reference is made to the second vector, it will be appreciated that the replication factor is optionally present in the cell other than following transfection with a first vector. For example a culture of cells that already express the replication factor may be obtainable from a third party.

In an embodiment of the invention described in detail below, the method comprises transfecting an ES cell with a first vector that expresses a viral replication factor, and thereafter transfecting the ES cell with a second vector that expresses a cDNA and is dependant upon presence of the viral replication factor for its extrachromosomal replication within the ES cell. The frequency of the first transfection step is generally low and may result in as few as 1 in 10^5 successful stable transfectants - this level of success is recognised as typical in this art. However, the second transfection has surprisingly and advantageously found to result in a significantly higher frequency of successful stable transfectant colonies being obtained. The second transfection can be carried out with a 1% or higher success rate, which represents a 100-fold improvement over the art.

One suitable viral replication factor for mouse cells, in particular mouse ES cells, is polyoma large T antigen, in which case the cell of step (a) expresses the polyoma large T antigen and the second vector comprises an origin of replication that binds the polyoma large T antigen, such as the polyoma replication origin, referred to as *Ori*. Another suitable viral replication factor for primate cells is based upon Epstein Barr virus, in which the primate cell of step (a) expresses the EBNA-1 antigen and the second vector comprises an origin of replication that binds EBNA-1, such as *OriP*. Viral replication factors are generally species - specific and so expression of DNA according to the invention is dependent upon choice of a replication factor appropriate to the cell. Polyoma large T has been described for use in mouse cells. EβNA-1 is suitable for human cells. Still further systems are optionally based on papilloma virus replication factors, for human cells, or SV40 virus large T antigen, for simian cells, and further suitable replication factors may also be selected from functional variants,

derivatives and analogues of these replication factors, such as temperature sensitive variants.

In use, the second vector is constructed according to standard techniques so as to contain a cDNA sequence or insert of interest operatively combined with a promoter to express the cDNA. The second vector is used to transfect an ES cell already expressing a replication factor and successful transfectants are recovered in which it is found that the second vector is stably maintained within the ES cell and expresses the cDNA with a more homogenous pattern than when prior art techniques are followed. Thus, the invention provides an advantageous method for expression of a cDNA in a cell.

In this context, "homogenous" in relation to expression of a cDNA in a colony of transfected ES cells is used to indicate that most cells, or a large proportion of cells, or preferably most cells, or more preferably substantially all cells, express the cDNA and "stable" is used to indicate that the cells continue to express the cDNA at a similar level and preferably at substantially the same level. In the examples carried out to date and described below, homogenous transfection is seen with the method of the invention to a greater extent than in the art methods. Also, in the examples carried out to date and described below the method results in more stable expression, meaning that expression does not alter over time. This has the advantage that study of the long term effects of a cDNA product is facilitated.

It is optional for the cell of step (a) first to be obtained or prepared by transfection of a cell by a first vector and for this then to be used for the starting cells for carrying out a plurality of separate transfections by second vectors containing different DNA inserts coding for different DNA products of interest. Following this procedure, the first transfection may be carried out with the level of success typically seen in conventional techniques and the ES cells obtained divided into separate colonies. The second transfections, introducing the DNA insert in the second vector, are then carried out with the higher levels of success typically seen in the methods of the invention.

In the case that the method comprises transfection with first and second vectors, it is preferable for the first vector to code for a selectable marker and for the second vector also to code for a selectable marker, though a different one. In a specific embodiment of the invention described below, the first vector codes for hygromycin resistance and the second codes for neomycin resistance. This allows selection of ES cells in which transfection by both first and second vectors has been successful.

It is a further embodiment of the invention for the method to comprise an additional transfection step with a third vector, wherein the third vector contains a cDNA, or is adapted to receive a cDNA, in operative combination with a promoter for expression of the cDNA, and extrachromosomal replication of the third vector is dependant upon presence within the ES cell of the replication factor. Transfection with the third vector is optionally at the same time as transfection with the second vector or subsequent thereto.

The second and third vectors preferably each comprise a selectable marker enabling selection of ES cells in which transfection has been successful. The respective selectable markers are preferably different if the method comprises transfection with both second and third vectors, and preferably different again from the selectable marker of the first vector.

It is a feature of particular embodiments of the invention that the second vector (and third or subsequent vectors if present) are not able to express the replication factor. In fact, in construction of the second vector from a vector comprising DNA encoding the replication factor it is preferable for that DNA to be largely or substantially completely deleted.

In a specific embodiment of the invention, the first vector is pMDG20neo and expresses polyoma large T antigen and the second vector comprises the natural target for polyoma large T antigen, namely *Ori*, expresses a cDNA of interest but does not express large T antigen. In use, the large T antigen is expressed by the first vector and binds to *Ori* of the second vector when it enters an ES cell, thus

enabling replication of the second vector and its maintenance within the ES cell in an extrachromosomal state. In successful transfectants, the vector remains extrachromosomal, and this is believed to render the vector relatively immune from effects seen when a vector is integrated into the host ES cell genome, which effect may include silencing of the cDNA resulting in unstable and heterogeneous expression.

An alternative to use of the first episomal vector is to introduce into the cell a construct that expresses the replication factor and integrates with the cell genome. The construct should therefore include a DNA sequence coding for the replication factor and means for selection of cells in which the construct has successfully integrated; one example is a construct that comprises cDNA coding for, in order, large T antigen - an internal ribosome entry site (IRES) - Bgeo. A culture of cells is then obtained by selecting for cells that express the selectable marker, such as in this case by selection in G418. Staining with Xgal is used to identify transfectant clones which show stable and homogenous expression. The construct preferably comprises a promoter that gives stable, low level expression in transfected cells, such as the HMGCoA promoter for ES cells. The cells obtained can then be subjected to transfection with the second and optionally third and subsequent vectors.

In another embodiment of the invention the second vector comprises an inducible promoter. Some types of differentiated cells, derived from ES cells, can only be obtained with any reliability if a particular differentiating factor is expressed after a prior event. One example is neurone formation which generally only occurs after aggregation of cells. Thus, using an inducible promoter, expression of DNA that codes for the factor that leads to neurone formation can be controlled until the ES cells have suitably aggregated. Interferon responsive promoters are some examples of inducible promoters. Alternatively, the cDNA is designed to be in a non-functional form and to be capable of being modified into a functional form at a later time. One possibility is for the cDNA to be disrupted for example by termination sequences which are flanked by target sites for a site specific recombinase, such as loxP sites, removable by Cre recombinase, or frt sites removable by Flp recombinase. Cre and

Flp can be fused to steroid hormone receptors in order to make their activity regulatable. After administration of steroid the Cre or Flp recombinase will translocate to the nucleus and there convert the cDNA into a functional form by excision of the disrupting sequence. It may also be desired to stop or inhibit or reduce replication of the second vector; the method optionally comprises using a site specific recombinase to present replication of the second vector. This can be achieved by deletion of a sequence in the vector to which the replication factor must bind in order for the vector to be replicated by the host cell.

The term DNA or cDNA is usually understood to refer to a DNA sequence that is transcribed into a mRNA that is translated into a polypeptide or protein. In the present invention the term is also intended to encompass any product of DNA expression. It thus includes DNA coding for an antisense RNA, or for an antisense ribozyme molecule.

The method of the invention is suitable for assaying effects of DNA expression, due to the stability and efficiency of expression achievable. Accordingly, the invention further relates to an assay for the effect of presence in a cell of any product of DNA expression - such as protein, polypeptide, antisense RNA, ribozyme RNA, transfer RNA or other. The method comprises steps (a) and (b) as described above wherein the second vector also contains a DNA coding for a selectable marker. The method further comprises selecting for cells that have been transfected with the second vector and maintaining the selected cells over a plurality of generations.

Step (a) may be carried out once and then steps (b) onwards repeated for different assays, and the method is of particular application to screening a cDNA library. Furthermore, two or more cDNAs can be expressed in the same cell to assay the effect of the combination of their respective expression products.

The invention also relates to a vector. Accordingly, the invention provides, in a second aspect, a vector for transfection of an ES cell, wherein:

(i) the vector contains a DNA, or is adapted to receive a DNA, in operative

combination with a promoter for expression of the DNA;

(ii) extrachromosomal replication of the vector is dependant upon presence within the ES cell of a replication factor; and

(iii) the vector does not express the replication factor.

The vector is characterized in preferred embodiments as described above in relation to the second vector of the first aspect of the invention.

It is an advantage of at least preferred embodiments of the invention that due to very high efficiency of stable secondary transfection (supertransfection) of cells, for example transfection of ES cells harbouring pMGD20neo with a second plasmid containing the polyoma replication origin (*Ori*) (8), that expression of DNA is stably and efficiently achieved from the second plasmid.

Another aspect of the present invention provides a method of screening for new DNAs that encode signal sequences and proteins that are transported to the cell surface. The invention according provides a method of investigating the properties of a DNA sequence comprising expressing in a cell a composite DNA including (a) the DNA sequence under investigation, linked to (b) a DNA coding for a cell active protein, wherein

activity of the cell active protein is dependant upon transport of the cell active protein to the cell surface, and

the DNA of (b) does not code for a polypeptide capable of directing transportation of the cell active protein to the cell surface.

This offers the advantage that where the DNA of interest does indeed code for a sequence that transports a polypeptide to the cell surface, whether that polypeptide remains there or is ultimately secreted, this will be apparent from observation that the cell active protein has had or is having its known effect. Thus the method offers a convenient means of identifying DNA sequences that will transport proteins to the cell

surface.

The method is suitably used for screening a library of DNAs to identify DNA sequences coding for signal polypeptide sequences that transport proteins to the cell surface. The cell active protein if transported to the cell surface may remain there or be secreted by the cell, and this distinction may be separately assayed, or example by examination of the make-up of the culture medium before and after the investigation.

One convenient way to obtain the DNA of (b) is by deleting or disabling, from a DNA encoding a cell surface or secreted protein, that portion of the DNA that codes for the polypeptide sequence responsible for transportation of the protein to the cell surface. The cell active protein is optionally a cell surface receptor and the DNA of (b) can thus encode a modified form of the receptor preprotein lacking a functional signal sequence. In a specific embodiment described below the IL-6 receptor is used as expression of the receptor in ES cells can be used to inhibit differentiation of the cells - a readily observable property of the cell active protein. Gross morphological or proliferative changes induced in the cell by the cell active protein are of course readily observed, though the invention is of application to any cell active protein whose activity, when it is transported to the cell surface and / or secreted, can be assayed.

A specific embodiment of this aspect of the invention comprises expressing the composite DNA by:

(a) (i) transfecting a cell with a first vector that expresses a replication factor; or

(ii) otherwise obtaining a cell that expresses the replication factor;

(b) transfecting the cell with a second vector, wherein

(i) the second vector contains the composite DNA in operative combination with a promoter for expression of the composite DNA;

(ii) the second vector also contains a DNA coding for a selectable

marker in operative combination with a promoter for expression of the selectable marker; and

(iii) extrachromosomal replication of the second vector is dependant upon presence within the cell of the replication factor;

(c) selecting for cells that have been transfected with the second vector; and

(d) maintaining the selected cells over a plurality of generations so as to assay the effect of expression of the composite DNA.

If many investigations are to be carried out it is preferred that step (a) is carried out once and the cells obtained are divided and used for a plurality of separate methods in which steps (b)-(d) are carried out a plurality of times with second vectors containing different DNA sequences. This offers the advantage that typically the first transfection step is of lower efficiency than the second, so the method avoids having to repeat the low efficiency step too often.

It is particularly preferred that the method is used for identification of a DNA coding for a cell surface or secreted protein, and using the method to screen a library of DNAs provides a means of carrying out the screen for discovery of such DNAs and investigation of their properties. More especially, the method is for discovery of hitherto unknown or uncharacterized cell surface or secreted proteins, or for location of the coding sequence of known proteins of this type.

This aspect of the invention optionally further incorporates in preferred embodiments features of transfection of cells described above in relation to other aspects of the present invention.

The invention enables development of a series of vectors which give highly efficient and robust expression of transgenes in cells. Cloned cDNAs of interest can rapidly be characterised using this system. It is also applicable to the discovery of novel regulatory molecules through functional expression screening of cDNA libraries.

Due to their pluripotent and proliferative character, key cellular processes such as viability, propagation, determination and differentiation, can be analyzed in transfected ES cells. The "supertransfection" system of the invention overcomes the limitations associated with conventional cDNA transfection and opens a powerful new route to gene discovery and characterisation in mammals.

Key features of the episomal supertransfection system, described according to the examples below, are that very high efficiencies of stable transfection are obtained and that cDNA expression is homogeneous, stable and reliably dictated by promoter strength. The increased efficiency of isolating stable transfectants is significant because it allows reliable detection of cDNAs whose expression results in cell death or differentiation. In addition a high transfection efficiency is generally advantageous for any high throughput assay system and is essential for functional cDNA library screening. The reliability of cDNA expression is critical for functional studies and the robust nature of expression from episomal vectors contrasts favourably with the variable and unstable expression observed in conventional ES cell transfectants.

Heterogeneous expression of integrated transgenes is not an artefact arising from use of bacterial *lacZ* as a reporter gene, firstly because similar observations have been made using mammalian thy-1 as a reporter in F9 cells, and secondly because ubiquitous expression of *lacZ* can readily be obtained following gene trap integrations (23,24). The expression pattern throughout the population cannot be determined by Northern blot but can only be revealed by *in situ* hybridization or use of a linked reporter gene such as IRES-*lacZ* (25) Heterogeneous expression, which previously occurred in the great majority of transfected clones following stable integration, gave unclear or misleading results on the phenotypic consequences of transgene expression.

The difference in expression pattern between conventional transfectants and episomal supertransfectants of the invention arises because an extrachromosomal copy of a transgene is not subject to alteration during the integration process nor to modification arising from the genomic sequences flanking an integration site. The so-called

"position effect" can modify both the level and pattern of transgene expression in stable transfectants. Furthermore, the expression of integrated transgenes is often suppressed over several generations in ES cell cultures. This silencing phenomenon contributes to the high backgrounds which can be obtained in double replacement type targeting strategies (26) . It has been observed in stable transfectants with different transgenes driven by viral promoters or minimal mammalian promoters such as the widely used human β-actin and mouse PGK-1 promoter elements. One hypothesis to explain this phenomenon is that transgenes may become targets of *de novo* methyltransferase in stem cells (27). Macleod et al. (28) reported that a methylation free locus could be generated in transgenic mice by introduction of the whole CpG island of the aprt promoter.

Whatever the molecular mechanism of silencing, it appears not to occur to episomally maintained transgenes in vectors of the invention. In addition, the level of expression obtained from vectors of the invention is reliably dictated by promoter strength and can predictably be varied over at least a 10-fold range by appropriate choice of promoter. Episomal constructs of the invention thus offer considerable advantages for functional expression studies in ES cells.

Functional cDNA expression cloning is a powerful method for direct isolation of important genes. The expression screening approach has often been employed to isolate cDNAs encoding surface and secreted molecules via transient expression, for example in COS cells. In a few cases EBV-based systems have also been applied to isolate intracellular regulatory genes via stable expression in the target cells (29-32) . The high efficiency of supertransfection in the polyoma system of the invention indicates that this approach could be applied to functional cloning in ES cells. Based on a transfection efficiency of 2.5%, a library of $5x10^5$ cDNA clones could be screened by electroporation of $2x10^7$ cells with 100μg DNA. For an effective library screen, the majority of transfectants should only take up a single plasmid. It is also advantageous if the cDNAs can readily be recovered in unrearranged form. Both of these conditions are satisfied by the episomal supertransfection system. By screening libraries prepared from undifferentiated ES cells it may be possible to

isolate cDNAs whose products mediate self-renewal. In this case direct selection can be applied for colony formation in the absence of LIF. For cDNAs whose products direct differentiation, however, it will be necessary either to screen pools through several rounds or to incorporate an inducible promoter into the episome.

Recently, several improved protocols for *in vitro* differentiation of ES cells have been reported, which promote efficient generation of, for example, haematopoietic cells (33) , neurons (34) or cardiomyocytes (35). The episomal expression strategy of the invention can be applied for gain-of-function assays and screens during these differentiation programmes. It can also be used for loss-of-function analyses via overexpression of anti-sense RNA or dominant-negative mutants. Combination of these differentiation systems with the episomal expression system will therefore provide powerful tools for analysing cell determination and differentiation events.

The invention is now described with reference to the accompanying drawings in which:

Fig. 1 shows the structure of the episomal expression vector pHPCAG;

Fig. 2 shows supertransfection efficiency of pHPCAG in MG1.19 ES cells;

Fig. 3 shows DNA hybridisation analysis of Hirt supernatants from supertransfectants;

Fig. 4 shows the effect of vector size on supertransfection efficiency;

Fig. 5 shows expression of β-galactosidase in MG1.19 transfectants;

Fig. 6 shows the restriction pattern of plasmid DNAs recovered from pHPCAG-*lacZ* supertransfectant clone;

Fig. 7 shows induction of differentiation by expression of STAT3F in MG 1.19 ES cells;

Fig. 8 shows co-supertransfection of STAT3F with wild type STAT expression vectors;

Fig. 9 shows linker sequences for use in an assay of the invention;

Fig. 10 shows DNA sequences coding for truncated and modified IL6R; and

Fig. 11 shows a vector for use in an assay of the invention.

In more detail:

Figure 1 shows the structure of the episomal expression vector pHPCAG. cDNAs can be introduced between two *Bst*XI sites using *Bst*XI adaptors. Abbreviations: ΔLT20: deleted polyoma large T expression cassette LT20; Pyori/enh: mouse polyoma virus replication origin and mouse polyoma mutant enhancer derived from F101 strain; SVpA: SV40 polyA addition signal; PGK*hph*pA: hygromycin B phosphotransferase gene expression cassette with mouse phosphoglycerokinase-1 (PGK) promoter and polyA addition signal; CAG: combined CAG expression unit; β-globinpA: rabbit β-globin polyA addition signal; SVori: SV40 replication origin; ColE1ori: ColE1 replication origin; *amp*: *E.coli* β-lactamase gene conferring resistance to ampicillin.

Figure 2 shows supertransfection efficiency of pHPCAG in MG1.19 ES cells.

(A) shows numbers of transfectant colonies per microgram of pHPCAG DNA. $5x10^6$ MG1.19 ES cells were supertransfected with the indicated amounts of supercoiled pHPCAG followed by selection with hygromycin B for 8 days. The resulting number of drug-resistant colonies were scored and efficiency per μg DNA calculated.

(B) shows total numbers of transfectant colonies plotted against total amount of plasmid DNA.

Figure 3 shows DNA hybridisation analysis of Hirt supernatants from supertransfectants. Hirt supernatants were prepared from $5x10^6$ parental MG1.19 cells and pooled pHPCAG supertransfectants. 1/20 of each sample was digested with either *Eco* RI or *Hind*III and analyzed by filter hybridisation using a 344bp *Sca* I-*Ssp*I fragment from pUC19 which is common to both pMGD20*neo* and pHPCAG..

Figure 4 shows the effect of vector size on supertransfection efficiency. 20μg of each of the supercoiled vectors pLT20Δ*Nde*I*hph* (4.7), pLT20Δ*Bst*X1*hph* (5.5), pLT20Δ*Alw*N1*hph* (5.6), pLT20Δ*Sac*I*hph* (5.9), ptkp (6.2), pSV40e/p (6.4), PGK*hph*ΔLT20 (6.5), pmPGKp (6.6), phBAp (6.6), pHPCAG (7.7), ptkp-*lacZ* (8.9);

pSV40e/p-*lacZ* (9.1), pmPGKp-*lacZ* (9.3), phBAp-*lacZ* (9.3), and pHPCAG-*lacZ* (10.4) were individually supertransfected into $5x10^6$ MG1.19 ES cells. The resulting numbers of hygromycin B resistant colonies were scored after 8 days. Transfection efficiencies are normalised relative PGKhphΔLT20.

Figure 5 shows expression of β-galactosidase in MG1.19 transfectants. Primary colonies were stained with Xgal after 8 days of selection.

(A) shows typical homogeneous staining pattern obtained following supertransfection with supercoiled pHPCAG-*lacZ*.

(B)shows heterogeneous staining pattern obtained in minority of clones following supertransfection with supercoiled pHPCAG-*lacZ*.

(C) shows heterogeneous staining pattern typically observed following electroporation of linearized pHPCAG-*lacZ* and stable integration.

(D) shows rare faint staining pattern obtained after supertransfection with supercoiled pHPCAG-*lacZ*.

Figure 6 shows the restriction pattern of plasmid DNAs recovered from pHPCAG-*lacZ* supertransfectant clone.

A supertransfectant MG1.19 clone carrying pHPCAG-*lacZ* was cultured for 60 days in the presence of hygromycin B. Hirt DNA was then prepared and electrotransformed into *E.coli* DH10B cells. Plasmid DNAs were recovered from transformants, digested with *EcoR*I, resolved by electrophoresis on 1.0% agarose gel and visualised by ethidium bromide staining. Expected fragment sizes: pMGD20neo, 4852bp and 2884bp; pHHPCAG-*lacZ*, 3697bp, 2810bp, 783bp and 397bp. Lane 1: size marker (1kb ladder:BRL); lane 2: control pMGD20; lane 3 : control pHPCAG-*lacZ*; lane 4: recovered pMGD20; lane 5,6: recovered pHPCAG-*lacZ*.

Figure 7 shows induction of differentiation by expression of STAT3F in MG 1.19 ES cells.

(A)shows proportion of differentiated colonies in LIF-supplemented medium resulting from supertransfection of STAT3, antisense STAT3 and STAT3F expression vectors. Colonies were fixed and stained with Leishman's reagent after 8 days selection and numbers of stem cell colonies and differentiated colonies scored.

(B) shows marker gene expression in STAT3F supertransfectants: Expression of marker genes in pools of MG1.19 cells supertransfected with STAT3 (lane 1), STAT3 antisense (lane 2) and STAT3F (lane 3) expression vectors. Total RNA was prepared after 8 days of selection in LIF-supplemented medium and 5μg aliquots analyzed by filter hybridisation with β-globin, Rex-1, H19 and G3PDH probes. The β-globin probe detects all transgene mRNA species generated from pHPCAG, including an alternatively spliced product from the antisense construct.

(C)shows photomicrographs of representative colonies 8 days after supertransfection with (i) STAT3, (ii) STAT3F, and (iii) empty expression vectors and selection in the presence of LIF, or, (iv) induction of differentiation by culture in the absence of LIF for 8 days.

Figure 8 shows co-supertransfection of STAT3F with wild type STAT expression vectors. Proportions of undifferentiated stem cell colonies generated after co-supertransfection of MG1.19 ES cells with 10μg pBPCAGGS-STAT3F plus 10μg pHPCAG vector containing stuffer (control), STAT3, STAT1 or STAT4 inserts. After 8 days selection with 80μg/ml of hygromycin B plus 20μg/ml of blasticidin S, colonies were fixed and stained with Leishman's reagent.

EXAMPLE 1

Materials and Methods

Vector constructions.

Standard recombinant DNA methods were used to construct all plasmids(10) . Plasmid pHPCAG (Fig 1) was constructed from pMGD20neo(8) . The PGK*neo*polyA sequence was replaced by a hygromycin resistance marker, PGK*hph*pA, and large T sequences were deleted (see Results). A *Sal*I-*Sca*I fragment containing the CAG expression unit, a *Bst*XI stuffer sequence, the polyA addition signal derived from the rabbit β-globin gene and an SV40 replication origin (11) was inserted. Coding sequences for β-galactosidase, LIF or interleukin-2 were introduced between the BstXI sites.

For construction of episomal expression vectors with alternative promoters, the *Sal*I-*Xba*I fragment containing the CAG expression unit in pHPCAG-*lacZ* was replaced with the 344 bp SV40 enhancer/promoter (SV40e/p), the 466 bp human β-actin promoter (hBA), the 502 bp mouse phosphoglycerate kinase promoter (mPGK) and the 90 bp HSV-tk minimal promoter (tk), resulting in pHPSV40e/p-*lacZ*, pHPhBA-*lacZ*, pHPmPGK-*lacZ* and pHPtk-*lacZ*, respectively.

Episomal vectors with alternative selection markers were constructed by replacing the PGK*hph*pA cassette in pHPCAG with the SV*bsr*pA cassette carrying the *E.coli* blasticidin S deaminase (*bsr*) gene derived from pSV2*bsr* (Waken Seiyaku) or the hCMVzeopA cassette carrying the *Streptoalloteichus* bleomycin resistant gene (*Sh ble*) derived from pZeoSV (Invitrogen) to generate pBPCAGGS and pZPCAGGS, respectively.

Cell culture and transfection.

MG1.19 ES cells are derivatives of the CCE line which stably maintain around 20 episomal copies of pMGDneo(8) . They were maintained on gelatin-coated plates in Glasgow modified Eagle's medium (GMEM, Gibco-BRL) supplemented with 10% fetal calf serum, 0.1 mM β-mercaptoethanol, non-essential amino acids, 200 μg/ml G418, and 100U/ml LIF produced in COS-7 cells(11,12) . For supertransfection, routinely,

$5x10^6$ MG1.19 cells were suspended in 800 μl of PBS, incubated with 20 μg of supercoiled vector DNA for 10 min on ice, and electroporated at 200V/960μF using a Bio-Rad gene pulser. Cells were transferred into gelatinized plates and allowed to recover overnight before addition of appropriate selection agent. Histochemical staining for β-galactosidase was carried out with 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-gal) (13) , and β-galactosidase activity was measured by incubation of cell extracts with o-nitrophenyl-β-D-galactopyranoside (ONPG). Differentiation was induced in monolayer culture as described (12) .

Analysis of episomal vectors in the supertransfectants.

Hirt supernatants were prepared as described (14) . For amplification of recovered episomal vectors, electrocompetent *E. coli* DH10B cells were transformed by electroporation at 2500V/25μF/200½.

Results

Construction of an episomal expression vector.

Polyoma-based plasmids have recently been reported to be competent for episomal propagation in ES cells (8) . The plasmid pMGD20neo contains a modified large T expression unit called LT20, the viral origin of replication (*Ori*), and the PGKneopA cassette as a selectable marker. This plasmid can be maintained as an extrachromosomal element in wild-type ES cells. It can be modified to include a cDNA expression unit (9) . However, the low frequency of conventional stable transfection of ES cells (Å 1 x 10^{-5}) remains a limiting feature. Furthermore, episomal propagation only occurs in 10-15% of primary transfectants (8,9) .

A second plasmid has been described which can be maintained as an episome only in ES cells which independently express the large T protein (8) . This plasmid, PGKhphΔLT20, contains LT20 with a large deletion in its coding sequence, *Ori*, and PGK*hph*pA as a selectable marker. When introduced into a cell line such as MG1.19, in which episomal maintenance of pMGDneo has already been established, the yield of hygromycin B resistant stable transfectants is extremely high. This phenomenon of supertransfection is presumed to arise from the pre-existence of large T protein in

the recipient cells.

In the studies reported below the modification and use of supertransfection vectors for cDNA expression is characterised.

Size of vector

PGK*hph*ΔLT20 retains part of the large T coding sequence. We made a series of deletions in the ΔLT20 sequence to minimize the vector size and thereby increase the capacity for inserts and reduce potential bias in the construction and screening of cDNA libraries. The supertransfection efficiency of four derivative plasmids was then compared in MG1.19 cells. All showed comparable supertransfection efficiency to PGK*hph*ΔLT20 (data not shown). The smallest, pLT20Δ*NdeIhph*, has a deletion of 2953 bp, yielding an episomal vector backbone of only 4.7kb.

Expression unit

Into this minimal episomal vector we introduced a cDNA expression unit. Transcriptional initiation signals are supplied by the CAG cassette(11) , which comprises the human cytomegalovirus immediate early enhancer, a 1kb fragment of the chicken β-actin gene (promoter, non-coding first exon and first intron), and a splice acceptor derived from the rabbit β-globin gene. This combination has been shown to direct strong expression of cDNAs in undifferentiated stem cells. The resulting expression vector, pHPCAG (Fig 1), contains the CAG sequences followed by the *BstXI* stuffer sequence derived from pCDM8 as a cDNA cloning site, and a polyA addition signal derived from the rabbit β-globin gene. In addition the plasmid contains the PGK*hph*pA (15) cassette for hygromycin selection of ES cell transfectants, the polyoma*Ori* with pyF101-derived mutant enhancer element (16) for stable episomal replication in cells expressing polyoma large T protein, and the β-lactamase (*amp*) gene and prokaryotic replication origin for amplification in *E. coli*. The SV40 *Ori* is also present to allow for transient episomal replication in mammalian host cells expressing SV40 largeT, such as COS cells (17) .

Characterization of supertransfection.

The parameters of supertransfection with pHPCAG and derivatives were investigated. First, $5x10^6$ MG1.19 cells were electroporated with various amount of supercoiled pHPCAG, selected in medium containing 80 μg/ml of hygromycin B for 8 days, and the number of stem cell colonies scored after Leishman's staining(12) . Although the highest efficiency per μg DNA was observed with minimum amounts (1-2 μg) of vector DNA (Fig. 2B), the total yield of hygromycin B resistant colonies increased with increasing amount of plasmid (Fig 2A). Saturation was not reached over the range of plasmid concentrations tested. With 100 μg plasmid DNA, 150,000 hygromycin B-resistant colonies were obtained, representing 3% of total treated cells. Disablement for episomal replication by linearisation of pHPCAG prior to electroporation reduced this transfection efficiency to less than 0.01%.

Next, increasing numbers of MG 1.19 cells were subjected to electroporation with 100 μg of pHPCAG DNA. Comparable stable transfection efficiencies in the range 3-6% were obtained with up to $2.5x10^7$ cells.

The copy number of pHPCAG in the supertransfectants was analyzed by preparation of Hirt supernatants followed by filter hybridisation. This analysis revealed that supertransfected cells carried approximately 20 copies each of pMGDneo and pHPCAG (Fig. 3).

These data demonstrate that the efficiency of supertransfection with pHPCAG is extremely high. However, episomal vectors can be limited in their capacity for inserts because increased size may cause inefficient replication or instability. To investigate this issue in the ES cell system, episomal vectors of different size were supertransfected into MG 1.19 cells. The numbers of supertransfectant colonies were scored and plotted against vector size (Fig. 4). These data indicate that there is a progressive reduction in transfection efficiency with increasing plasmid size. In particular, the largest plasmid tested, a derivative of pHPCAG with a 3kb *lacZ* insert (total size 10.4kb) showed a 50% reduction in colony number. However, that this may not be due entirely to the size of the plasmid because the very high levels of β-

galactosidase expression may exert some toxic effects (see below).

***lacZ* expression in supertransfectants.**

To evaluate the level and pattern of expression of transgenes from pHPCAG, the *E.coli* β-galactosidase (*lacZ*) gene was introduced into this vector. The resulting vector, pHPCAG-*lacZ*, was introduced into MG1.19 cells and supertransfectants isolated by selection with 80 μg/ml of hygromycin B for 8 days. The number of colonies isolated was 50% of the number obtained in a parallel supertransfection with pHPCAG (see above). The colonies were smaller and many of the cells showed an abnormal spindle-shaped morphology. These effects were not observed with several other inserts in pHPCAG and are suggestive of a toxic effect of the high level *lacZ* expression. The primary supertransfectants were stained with X-gal and the staining pattern examined under phase-contrast microscopy. Staining was detectable after 5 minutes incubation and was intense by 1 hour. This level of β-galactosidase activity is significantly higher than we have observed from a variety of integrated expression constructs.

Approximately 80% of supertransfectant colonies showed ubiquitous expression (>90% cell positive) as shown in Fig.5-A (i). Of the remainder, 15% showed heterogeneous expression (Fig.5-A (ii)), and 5% showed little or no staining (Fig.5-A (iv)). The latter two classes are likely to arise as a result of vector integration which occurs in up to 20% of supertransfectants (8). In transfectants derived by electroporation of linearized pHPCAG-*lacZ* into MG1.19 cells (which results in vector integration in the majority of clones), only 15 % of colonies showed homogeneous staining whereas 70% of colonies stained heterogeneously (Fig.5-A (iii)), and 15% showed no expression.

Analysis of expanded clones from each class of transfectant established that this difference in expression characteristics was stable. Twelve of 13 expanded supertransfectants expressed *lacZ* homogeneously. In contrast, only 4 out of 24 clones derived using linearized vector showed homogeneous expression. This is consistent with our previous observations on integrated expression constructs in ES

cells. In fact the CAG unit gives a significantly higher frequency of colonies which show stable ubiquitous expression than other promoters we have examined.

The difference in staining pattern between episomally maintained and integrated vectors indicates that the former escape modifying influences arising from integration and reliably give full activity of the expression unit.

Comparison of expression with various promoters on episomal vector.

An ability reliably to generate predetermined levels of expression would be a important attribute for a transgene expression system. The previous observations suggested that episomal vectors offered potential to achieve unmodified expression. Various promoters with different strengths in undifferentiated stem cells were therefore introduced into the episomal vector by replacing the CAG expression unit of pHPCAG-*lacZ*. Expression of the *lacZ* reporter was then assayed in both transient and stable supertransfectants (Table 1). The relative ratio of β-galactosidase activity obtained from the SV40 enhancer/promoter complex, the human β-actin promoter, the mouse PGK-1 promoter and the HSV-tk minimal promoter in transient transfectant was retained in stable supertransfectants. The CAG expression unit showed strongest activity in the tested constructs in both transient and stable transfectants. In this case, however, the relative ratio in transient transfectants, 19 times higher than SV40, was significantly reduced in stable transfectants. This may arise from an elimination of strong expressants due to a toxic effect of high *lacZ* expression (see above). A reduced number of supertransfectants and smaller size of colonies was observed only with the CAG vector.

Stability of supertransfected episomal expression vector during long-term culture and differentiation of host cells.

A critical limitation of previously described episomal vectors is their instability during long-term culture. Many episomal vectors undergo integration into the host genome after long-term culture, resulting in a reduction in expression and inability to recover transgenes by preparing Hirt supernatants. To test the stability of the supertransfection system, four pHPCAG-*lacZ* supertransfectant clones were cultured

for 60 days (approximately 90 generations) under continuous selection with 80 μg/ml of hygromycin B. Three of the four clones maintained relatively constant levels of β-galactosidase activity determined by ONPG assay and uniform expression as revealed by Xgal staining. The fourth clone showed unstable and variegated expression, as commonly observed on vector integration. Hirt supernatants were prepared from one of the stably expressing clones at the end of the 60 day culture period. Filter hybridization analysis of the Hirt DNA indicated that the ES cells carried approximately 20 copies of pMGD20 and 5 copies of pHPCAG-*lacZ* per cell (data not shown). The lower copy number of pHPCAG-*lacZ* may be due to its larger size and/or the toxic effect of strong *lacZ* expression. The Hirt DNA was transformed into *E.coli* for further analysis. Of the bacterial transformants, 20% carried pHPCAG-*lacZ* and the remainder carried pMGD*neo*20, in good agreement with the hybridization data. Restriction mapping showed no evidence of rearrangement in either plasmid (Figure 6).

In the experiment above, cells were maintained under selection with hygromycin B. In the absence of selection pressure, supertransfectant clones lost expression of β-galactosidase over several passages in culture. This might indicate an intrinsic instability of supertransfected episomal vectors. However, it could also reflect a selective disadvantage for ES cells which express high levels of β-galactosidase. It is noteworthy in this regard that the primary episome, pMGD20neo, is stable in the absence of selection(8).

Stability of expression from pHPCAG-*lacZ* during the *in vitro* differentiation of ES cells was also analyzed. Differentiation was induced in three ways: withdrawal of LIF; exposure to retinoic acid; and treatment with 3-methoxybenzamide(18). After 6 days the differentiated progeny stained ubiquitously in all three cases (data not shown).

These data indicate that supertransfected episomal vectors can be maintained in an extrachromosomal state and direct strong expression of transgenes during long-term self-renewal and differentiation *in vitro*.

Production and secretion of the cytokine LIF from an episomal ES cell expression vector.

The pHPCAG-*lacZ* plasmid can efficiently direct strong and homogeneous expression of the cytoplasmic *lacZ* reporter gene. We next investigated expression of a secreted molecule, the cytokine LIF. LIF is an essential supplement to ES cell culture medium because it inhibits differentiation of the stem cells (19,20) . Expression of LIF can readily be assayed by formation of stem cell colonies in media lacking the cytokine.

Episomal vectors for expression of another cytokine, interleukin-2 (which has no effect on ES cell phenotype), and for LIF were electroporated in parallel into MG1.19 cells. The cells were seeded at low density ($1.5x10^4$ and $5x10^3$ cells per 90mm plate) to avoid the rescue effect which arises from the production of LIF by differentiated ES cell progeny (21) , and cultured with 80 μg/ml of hygromycin B for 8 days. pHPCAG-*il2* generated large numbers of stem cell colonies in medium supplemented with LIF, but none in the absence of LIF. pHPCAG-*lif* in contrast produced comparable numbers of healthy stem cell colonies in both the presence and absence of exogenous LIF (Table 2). These colonies could be expanded and propagated without LIF-supplementation of the medium. These data confirm previous observations that increased autocrine expression of LIF renders ES cells factor-independent (22) and establish that secreted proteins are produced efficiently and stably by this episomal expression system.

Co-supertransfection of episomal vectors.

Introduction of two or more different transgenes into cells is often required for analysis of protein interactions and/or co-operative function. The poor efficiency of homogeneous expression in conventional transfectants is a major obstacle for such investigations in ES cells. To test the possibility that the episomal approach could be applied to co-express multiple cDNAs, we constructed episomal expression vectors with different selection markers. Co-supertransfection of episomal vectors was then assessed.

The basic episomal expression vector pHPCAG carries the hygromycin

phosphotransferase gene driven by mouse PGK-1 promoter (PGK*hph*pA). We prepared episomal vectors which carry the zeocin-resistance gene driven by the human cytomegalovirus immediate-early promoter (pZPCAG), or the blasticidin S-resistance gene driven by the SV40 enhancer/promoter (pBPCAG) by substitution of the PGK*hph*pA cassette in pHPCAG. These vectors were supertransfected into MG1.19 cells followed by 8 days selection with the appropriate antibiotic. Comparison of the numbers of resulting drug-resistant colonies (Table 3) revealed that these selection systems are slightly less efficient than hygromycin B selection but nonetheless enable large numbers of supertransfectants to be isolated.

ES cells harbouring two different episomal vectors can be isolated by repeated supertransfection. Supertransfectants carrying pHPCAG can be transfected again with pBPCAG or pZPCAG, with comparable efficiency to the original supertransfection into MG1.19 ES cells (data not shown). This should allow establishment of efficient screens for assaying functional interactions between gene products.

The effects of co-electoporation of supertransfection vectors were also investigated. pHPCAG (10 μg) and pBPCAG (10 μg) were co-electroporated into $5x10^6$ MG1.19 cells. Cells were selected in hygromycin B or blasticidin S only, or both, for 8 days and the number of drug-resistant colonies scored in each case. The numbers of hygromycin or blasticidin S single-resistant colonies were 39,000 and 13,000, respectively, while the number of double-resistant colonies was 1,200. Thus the apparent efficiency of incorporation of both plasmids was less than 10%. Similar results were obtained on co-supertransfection of pHPCAG and pZPCAG (not shown). These data suggest that the majority of supertransfectants incorporate only one plasmid under these electroporation conditions. This is significant for application of the episomal system to functional cDNA library screening.

EXAMPLE 2

The effects of overexpression of a large number of transgenes in ES cells were investigated by construction of vectors based on pHPCAG and including a DNA insert coding for the transgene being investigated. 5 x 10^6 ES MG1.19 cells were

supertransfected with 20 μg of expression vectors and selected with 80 μg/ml of hygromycin B for 8 days. The numbers of drug-resistant colonies were counted and normalised relative to numbers obtained with empty vector. The results are shown in Table 4.

EXAMPLE 3

Inhibition of STAT3 activation blocks self-renewal and promotes differentiation

To assess directly the requirement for STAT3 activation in ES cell self-renewal, we exploited a dominant interfering mutant form of STAT3, STAT3F. In this mutant (Minami *et al.*, 1996), the tyrosine residue at amino acid position 705 is mutated to phenylalanine. Phosphorylation of Tyr705 is required for dimerization and nuclear translocation. When expressed at high level, STAT3F has been shown to block the activation of endogenous STAT3 in various cell types, possibly by titrating out receptor docking sites (Fukada *et al.*, 1996; Minami *et al.*, 1996; Nakajima *et al.*, 1996; Bonni *et al.*, 1997; Ihara *et al.*, 1997).

Using conventional transfection approaches we were unable to recover ES cell transfectants showing stable high level expression of STAT3F. In parallel experiments, however, transfection of the LIF-independent embryonal carcinoma cell line P19 yielded multiple expressing clones. This suggested that blockade of STAT3 activation in ES cells specifically resulted in cell death, growth arrest or differentiation. The transfection and expression strategy of the invention was therefore adopted to enable characterisation of the consequences of STAT3F expression.

The STAT3F mutant cDNA was introduced into the supertransfection vector pHPCAG. The wild type STAT3 coding sequence was also introduced, in both sense and antisense orientations. The three constructs were electroporated into MG1.19 cells which harbour a large T expression plasmid and can be supertransfected with constructs containing the polyoma origin (Gassmann *et al.*, 1995). Supertransfectants were isolated by selection in hygromycin B for 8 days in the presence of LIF. Colonies were fixed, stained with Leishman's reagent, counted, and scored for the presence of stem cells and differentiated cells. More than 95% of colonies obtained

following supertransfection with control or wild type STAT3 vector were stem cell colonies (Figure 7A). A modest increase in the proportion of differentiated colonies was obtained with the antisense construct. The STAT3F vector, however, yielded predominantly differentiated colonies. A decrease in total number of colonies was also observed after supertransfection with STAT3F. This may reflect an early onset of differentiation which would produce very small clones that would not be scored. Alternatively, very high levels of STAT3F expression may also be toxic, though this has not been reported in other cell types. Morphologically, the differentiated STAT3F colonies closely resembled the differentiated colonies generated on culture of ES cells in the absence of LIF (Figure 7C). Various other cDNAs have been expressed in ES cells using this system, with little or no effect on differentiation (data not shown). This suggested that the effect on differentiation was specifically attributable to expression of STAT3F.

The differentiation induced by expression of STAT3F was examined further by expression analysis of the marker genes *rex1* and *H19*. Rex-1 mRNA, which is specifically expressed in undifferentiated stem cells, was down regulated in STAT3F supertransfectants. In contrast, H19 RNA which is found at low levels in stem cells but is upregulated during differentiation, was increased (Figure 7B). A similar pattern of gene regulation is observed during differentiation of ES cells induced by withdrawal of LIF. These data confirm that the morphological differentiation triggered by STAT3F is accompanied by reprogramming of gene expression.

STAT3F was also expressed from the mouse phosphoglycerate kinase (*pgk-1*) promoter in the episomal vector pHPPGK. This vector gives at least 10-fold lower expression than pHPCAG (data not shown). In this case, there was no significant effect on either colony number or differentiation status of MG1.19 supertransfectants. A critical level of expression of the dominant interfering mutant therefore appears necessary to block self-renewal.

Effect of STAT3F on self-renewal is suppressed by co-expression of STAT3

To test whether the induction of differentiation by expression of STAT3F was due to an inhibition of endogenous STAT3 activity, we attempted to rescue the stem cell phenotype by co-expression of wild type STAT3 and also of STAT1 and STAT4. A STAT3F expression vector carrying a blasticidin resistance marker was co-supertransfected into MG1.19 cells with episomal constructs for expression of wild type STATs and hygromycin resistance. Co-supertransfectants were isolated in medium containing both 20μg/ml of blasticidin S and 80μg/ml of hygromycin B. The numbers of stem cell and differentiated colonies were scored after 8 days. As shown in Figure 8, only co-expression of wild type STAT3 restored self-renewal in the presence of STAT3F. Transfection with STAT1 or STAT4 constructs alone had no effect on self-renewal in the absence of STAT3F (not shown) and did not alter differentiation induced by STAT3F. In the case of supertransfection with the CAG promoter STAT1 construct, the total number of colonies (stem + differentiated) recovered was reduced but the relative proportion of stem cell colonies versus differentiated cells was unaltered. This occurred in both the presence and absence of co-expression of STAT3F, and suggests that high level expression of STAT1 may be toxic to ES cells. By using the mouse PGK-1 promoter to drive lower levels of expression comparable numbers of colonies were recovered on transfection with the STAT1 as with the other constructs. In this case, again only the STAT3 construct showed any restoration of stem cell colonies, although to a lower degree than with the high expression CAG vector (not shown). These data indicate that STAT3 has a specific function in ES cells which cannot be compensated by STAT1 or STAT4.

EXAMPLE 4

The invention is also used in a strategy for direct selection of genes that code for secreted and cell surface proteins. In one example of this strategy, the basic cloning vector is a truncated form of IL6R that lacks a signal sequence. This vector is described in detail below and shown in Fig. 11. If this truncated IL6R is expressed in ES cells, it is not exported to the cell surface and these cells differentiate when cultured in IL6. However, if the IL6R signal sequence is reconstituted by a signal sequence provided by a cDNA fragments cloned in frame at the 5' end of the

truncated IL6R, the chimaeric receptor is expressed on the surface of ES cells. ES cells containing such chimaeric receptors are thus maintained as undifferentiated colonies when cultured in IL6.

Libraries of short, 5' cDNA fragments are produced and cloned into a truncated and modified IL6R-based expression vector. ES cells transformed with such libraries express cDNA:IL6R fusion proteins. However, only cDNAs that encode signal sequences confer IL6 responsiveness on ES cells. These cDNAs alone give rise to undifferentiated, proliferating ES cell clones. This strategy therefore provides a direct selection for cDNAs encoding secreted and cell surface proteins.

The chimaeric IL6R is expressed in the episomal expression system described above (or a derivative thereof). This allows drug selection for episomally transformed cells and high level expression of cloned DNA.

To further refine the selection system, ES cells are modified with two targeted mutations:

a) A selectable marker gene, for example the blasticidin resistance gene, is introduced into the OCT-4 locus by standard targeting techniques. Since Oct-4 is expressed in undifferentiated ES cells, the blasticidin resistance gene will be expressed only by undifferentiated colonies. Blasticidin selection therefore is used to decrease background growth by ensuring rapid deletion of differentiating, Oct-4 negative, ES cells.

b) Since ES cells can produce LIF as an autocrine growth factor, ES cells are used in which both copies of the LIFR gene have been disrupted by gene targeting. This eliminates the possibility of LIF-dependent, false positive colonies that might otherwise persist throughout selection in IL6.

Details of vector construction:

1). IL6R was cloned into the episomal vector pCAGIP or a derivative (pCAGIPXN,

i.e. pCAGIP with a destroyed NotI site). pCAGIP contains an internal ribosome entry site (IRES) and a puromycin resistance gene downstream of its multiple cloning site, resulting in stoichiometric production of cDNA:IL6R fusion proteins in transfected cells under puromycin selection. IL6R in pCAGIP provides a positive control (IL6-responsive functional protein on the cell surface), and the basis of the new vector.

2). To construct the cloning vector, IL6R cDNA was truncated by cleavage with BssHII at nucleotide number 92. This deleted the initiator ATG and sequences encoding the signal sequence.

3). To minimise potential steric interference by cloned proteins with IL6 binding and IL6R function, DNA encoding a synthetic flexible linker peptide was then added to the 5' end of the truncated IL6R. Two alternative linkers have been used: gly gly gly gly ser gly gly gly gly ser and a linker containing the FLAG epitope, gly ser ASP TYR LYS ASP ASP ASP ASP LYS (FLAG epitope in upper case). The sequence of these linkers is shown in Fig. 9. In each case, the linker sequence has been cloned in frame with IL6R and has two unique cloning sites (XhoI and NotI) at its 5' end, allowing the introduction of cDNA libraries, or specific cloned sequences, in a directional manner. The FLAG epitope is recognised by a commercially available monoclonal antibody (M2; available from IBI/Kodak) regardless of its position within a fusion protein, and will thus allow the expression levels of surface protein to be measured directly by immunocytochemistry.

4). Vectors containing each of these linkers and an upstream signal sequence are tested for relative expression level and IL6R-function, as detailed below.

To test the utility of these vectors for selecting proteins expressed at the cell surface, a number of known signal sequences are cloned into each vector. These are tested for surface expression and IL6R function. Signal sequences include those from rat CD4 (a protein with extracellular Ig domains), mouse sek (a receptor tyrosine kinase, with no extracellular Ig domains) and mouse sonic hedgehog (a secreted factor).

ES cells are transfected with vectors bearing candidate signal sequences by lipofection or electroporation, followed by puromycin selection for transfected cells. After overnight growth in the presence of LIF, to maintain the undifferentiated state and proliferation, transfected cells are split into three groups and treated with either 1) LIF, 2) IL6 or 3) neither growth factor. Only cells bearing IL6R brought to the cell surface by a fused signal peptide will proliferate in the presence of IL6. Positive controls include ES cells transfected with wild-type IL6R grown in the absence of LIF and the presence of IL6. Negative controls include empty vector (i.e truncated IL6R with no 5' insert) grown in the presence of IL6. To determine whether fusion proteins N-terminal to IL6R block signalling (by steric hindrance), the proportion of such cells that express surface protein but fail to proliferate in response to IL6 is deduced by comparing the number of cells expressing the FLAG epitope with the number that give rise to colonies.

Vectors defined by this assay are then used in cDNA library screens. Preferably, sequences corresponding to 5' ends of cDNAs are generated from full length cDNA libraries and directionally cloned in the screening vector.

In the above description scientific publications are referred to under the following reference numbers:

1. Smith, A. G. (1992) *Seminars in Cell Biology*, **3**, 385-399.
2. Evans, M. J. and Kaufman, M. (1981) *Nature*, **292**, 154-156.
3. Martin, G. R. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 7634-7638.
4. Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W. and Kemler, R. (1985) *J.Embryol.Exp.Morphol.*, **87**, 27-45.
5. Weiss, M. J. and Orkin, S. H. (1996) *J. Clin. Invest.*, **97**, 591-595.
6. Bradley, A., Evans, M. J., Kaufman, M. H. and Robertson, E. (1984) *Nature*, **309**, 255-256.
7. Beddington, R. S. P. and Robertson, E. J. (1989) *Development*, **105**, 733-737.
8. Gassmann, M., Donoho, G., Berg, P. (1995) *Proc.Natl.Acad.Sci.USA*, **92**, 1292-1296.
9. Camenisch, G., Gruber, M., Donoho, G., Van Sloun, P., Wenger, R. and

Gassmann, M. (1996) *Nucleic Acids Research*, **24**, 3707-3713.

10. Sambrook, J., Fritsch, E. F. and Maniatis, T. (eds.) (1989) Molecular Cloning: A Laboratory Manual. 2nd ed Ed. 3 vols. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
11. Niwa, H., Yamamura, K.-I., Miyazaki, J.-I. (1991) *Gene*, **108**, 193-200.
12. Smith, A. G. (1991) *J. Tiss. Cult. Meth.*, **13**, 89-94.
13. Beddington, R. S. P., Morgenstern, J., Land, H. and Hogan, A. (1989) *Development*, **106**, 37-46.
14. Hirt, B. J. (1969) *J. Mol. Biol.*, **26**, 141-144.
15. te Riele, H., Maandag, E. R., Clarke, A., Hooper, M. and Berns, A. (1990) *Nature*, **348**, 649-651.
16. Fujimura, F. K., Deininger, P. L., Friedmann, T. and Linney, E. (1981) *Cell*, **23**, 809-814.
17. Tsui, L. C., Breitman, M. L., Siminovitch, L. and et al. (1982) *Cell*, **309**, 499-508.
18. Smith, A. G. and Rathjen, P. D. (1991) *Sem.Dev.Biol.*, **2**, 317-327.
19. Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M. and Rogers, D. (1988) *Nature*, **336**, 688-690.
20. Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A. and Gough, N. M. (1988) *Nature*, **336**, 684-687.
21. Rathjen, P. D., Nichols, J., Toth, S., Edwards, D. R., Heath, J. K. and Smith, A. G. (1990) *Genes Dev.*, 4, 2308-2318.
22. Conquet, F., Peyrieras, N., Tiret, L and Brulet, P. (1992) *PNAS*, **89**, 8195-8199.
23. Skarnes, W. C., Auerbach, B. A. and Joyner, A. L. (1992) *Genes Dev.*, **6**, 903-918.
24. Friedrich, G. and Soriano, P. (1991) *Genes Dev.*, **5**, 1513-1523.
25. Mountford, P., Smith, A.G. (1995) *Trend Genet.*, **11**, 179-184.
26. Askew, G. R., Doetschman, T. and Lingrel, J. B. (1993) *MCB*, **13**, 4115-4124.
27. Tucker, K. L., Beard, C., Dausman, J., Jackson-Grusby, L., Laird, P.W., Lei, H., Li, E., Jaenish, R. (1996) *Genes & Dev.*, **10**, 1008-1020.

28. Macleod, D., Charlton, J., Mullins, J., Bird, A. (1994) *Genes Dev.*, 8, 2282-2292.

29. Strathdee, C. A., Gavish, H., Shannon, W. R. and Buchwald, M. (1992) *Nature*, **356**, 763-767.

30. Legerski, R. and Peterson, C. (1992) *Nature*, **359**, 70-73.

31. Hayakawa, H., Koike, G., Sekiguchi, M. (1990) *J. Mol. Biol.*, **213**, 739-747.

32. Sasaki, K., Watanabe, E., Kawashima, K., Sekine, S., Dohi, T., Oshima, M., Hanai, N., Nishi, T., Hasegawa, M. (1993) *J. Biol. Chem.*, **268**, 22782-22787.

33. Nakano, T., Kodama, H., Honjo, T. (1994) *Science*, **265**, 1098-1101.

34. Bain, G., Kitchens, D., Yao, M., Huettner, J.E., Gottlieb, D.I. (1995) *Dev. Biol.*, **168**, 342-357.

35. Rohwedel, J., Maltsev, V., Bober, E., Arnold, H.H., Hescheler, J., Wobus, A.M. (1994) *Dev. Biol.*, **164**, 87-101.

We have thus described the development of an optimised transfection and expression system which will enable high throughput functional screening of cDNAs in pluripotential mouse embryonic stem (ES) cells and differentiated derivatives. The strategy is based on extrachromosomal vector replication driven by expression of polyoma large T protein. When a vector containing a polyoma origin of replication is introduced into an ES cell line that harbours polyoma large T antigen, a high frequency of stable secondary transfection results. This process is referred to as supertransfection. Supertransfected plasmids can be maintained episomally during long-term culture and during differentiation *in vitro*. Expression of a β-galactosidase reporter from an episomal vector is both ubiquitous and stable, in contrast to the variegated and unstable expression usually observed after cDNA integration into the ES cell genome. Moreover, in the absence of integration, promoter strength is predictable and a range of expression levels can reliably be achieved by using different elements. We also show that episomal vectors can be used for efficient expression of both cytosolic and secreted proteins. These features should make this system invaluable for functional analyses of defined cDNAs and for direct expression screening of cDNA pools or libraries in ES cells.

Table 1. Comparison of β-galactosidase activities directed by various promoters in transient and stable supertransfectants.

Promoter	Relative β-gal activity	
	transient	stable
SV40 e/p	1.0	1.0
hβAp	1.1	0.7
mPGKp	0.5	0.5
TKp	0.1	0.1
CAG	19.0	1.8

$5x10^6$ MG1.19 ES cells were supertransfected with 20μg of vector DNAs. After 3 days culture for transient expression assay or 8 days selection with hygromycin B for stable expression assay, the β-galactosidase activity generated by these constructs was measured by ONPG assay. Results are normalised relative to activity generated by the SV40e/p construct. See 'Materials and methods' for construction details of vectors.

Table 2. Supertransfection of LIF and IL-2 expression vectors into MG1.19 ES cells.

Vector	LIF in medium	No. of hyg^r stem cell colonies
pHPCAG-*lif*	+	42,000
pHPCAG-*lif*	-	38,000
pHPCAG-*il2*	+	48,000
pHPCAG-*il2*	-	0

$5x10^6$ MG1.19 ES cells were supertransfected with 20μg of vector DNAs. After 8 days selection with 80μg/ml of hygromycin B in the presence or absence of LIF, the number of stem cell colonies were scored.

Table 3. Efficiency of supertransfection of vectors with various selection markers.

Selection marker	Drug for selection (μg/ml)	No. of resistant colonies
PGK*hph*pA	hygromycin B (80)	50,000
SV*bsr*pA	blasticidin S (4)	12,600
hCMV*zeo*pA	zeocin (20)	20,600

$5x10^6$ MG1.19 ES cells were supertransfected with 20μg of vector DNAs of episomal vectors, pBPCAG and pZPCAG, which carry *bsr* and *zeo* resistance genes respectively. After 8 days selection with the appropriate drug, the number of drug-resistant stem cell colonies were scored.

Table 4. Effects of overexpression of transgenes in ES cells using pHPCAG.

cDNA	Relative number of hygro^R colonies	Colony Size and Morphology
None	1.00	Normal
lacZ	0.64	small
DIA/LIF	0.87	slightly small
IL-2	0.92	slightly small
Rex-1	0.88	Normal
Fgf-2	0.65	Normal
Fgf-4	0.82	Normal
Fgf-5	0.41	Normal
Oct-1	0.17	small
Oct-2	0.65	slightly small
Oct-3/4	0.61	differentiated
Oct-6	0.03	some differentiation
c-jun	0.47	small
E1A	0.08	differentiated
Jak2 K/E	0.75	Normal
bcl-2	0.28	small, spindle morphology
MAPKP	1.38	Normal
RXRα	0.20	some differentiation
RXRβ	0.63	Normal
RXRγ	0.91	Normal
COUP-TF1	0.40	some differentiation
HNF-4	0.05	Normal
Stat1	0.10	small
Stat3	0.52	Normal
Stat4	0.16	Normal
Stat3DON*	0.14	differentiated

$5x10^6$ ES MG1.19 cells were supertransfected with 20 µg of expression vectors and selected with 80 µg/ml of hygromycin B for 8 days. The numbers of drug-resistant colonies were counted and normalised relative to numbers obtained with empty vector.

*Stat3DON is the dominant interfering mutant form of Stat3 described by Akira *et al.* (1996).

Claims

1. A method of expressing a DNA in a cell, comprising:

(a) (i) transfecting the cell with a first vector that expresses a replication factor; or

(ii) otherwise obtaining a cell that expresses or will express the replication factor;

and

(b) transfecting the cell with a second vector, wherein

(i) the second vector contains a DNA, or is adapted to receive a DNA, in operative combination with a promoter for expression of the DNA; and

(ii) extrachromosomal replication of the second vector is dependant upon presence within the cell of the replication factor.

2. A method according to Claim 1 wherein the replication factor is a viral replication factor.

3. A method according to claim 1 or 2 wherein the viral replication factor is selected from polyoma large T antigen, EBNA-1 antigen, papilloma virus replication factors, SV40 large T antigen and functional variants, analogues and derivatives thereof appropriate to the cell species.

4. A method according to any of claims 1-3 wherein the second vector does not express the replication factor.

5. A method according to any of claims 1-4 wherein the second vector expresses a selectable marker.

6. A method according to any of claims 1-5 further comprising transfecting the cell with a third vector, wherein the third vector contains a DNA, or is adapted to receive a DNA, in operative combination with a promoter for expression of the DNA, and replication of the third vector is dependant upon presence within

the cell of the replication factor.

7. A method according to Claim 6 wherein the third vector expresses a selectable marker, which selectable marker is different to that expressed by the second vector.

8. A method according to any preceding claim wherein the cell is a mammalian cell or an avian cell.

9. A method according to any preceding claim wherein the cell is an embryonic cell.

10. A method according to Claim 9 wherein the cell is an ES, EC or EG cell.

11. A method according to any preceding claim for transfection of an ES cell wherein the ES cell of step (a) expresses polyoma large T antigen and the second vector comprises a natural target for polyoma large T antigen, such as *Ori* or functional variants thereof adapted to bind to polyoma large T antigen.

12. A method according to any preceding claim wherein the DNA codes for a polypeptide or protein.

13. A method according to any of Claims 1-11 wherein the DNA codes for an antisense RNA.

14. A method according to any preceding claims wherein the promoter is inducible.

15. A method according to any preceding claim wherein transcription of the DNA can be activated by a site specific recombinase.

16. A method according to any preceding claim wherein replication of the second vector can be prevented by a site specific recombinase.

17. A vector for transfection of a cell, wherein:

(i) the vector contains a DNA, or is adapted to receive a DNA, in operative combination with a promoter for expression of the DNA;

(ii) extrachromosomal replication of the vector is dependant upon presence within the cell of a replication factor; and

(iii) the vector does not express the replication factor.

18. A vector according to Claim 17 wherein the replication factor is a viral replication factor.

19. A vector according to Claim 17 or 18 wherein the viral replication factor is selected from polyoma large T antigen, EBNA-1 antigen, papilloma virus replication factors, SV40 large T antigen and functional variants, analogues and derivatives thereof.

20. A vector according to any of Claims 17 to 19 wherein the vector is substantially free of DNA coding for the replication factor or any part thereof.

21. A vector according to any of Claims 17 to 20 for transfection of mammalian or avian cells.

22. A vector according to any of Claims 17 to 21 for transfection of ES cells.

23. A vector according to Claim 22 comprising a natural target for polyoma large T antigen, such as *Ori* or functional variants thereof adapted to bind to polyoma large T antigen.

24. A vector according to any of Claims 17-23 wherein the DNA codes for a polypeptide or protein.

25. A vector according to any of Claims 17-23 wherein the DNA codes for an antisense DNA.

26. A vector according to any of Claims 17-25 wherein the promoter is inducible.

27. A vector according to any of Claims 17 to 26 wherein the vector comprises a sequence coding for a selectable marker.

28. Use of a vector according to any of Claims 17-27 for expression of a DNA sequence within a cell.

29. A cell transfected with a first vector that expresses a replication factor and with a second vector according to any of Claims 17 to 27.

30. A mammalian cell according to Claim 29.

31. An embryonic cell according to Claim 29.

32. A cell selected from an ES, EC or EG cell according to any of Claims 29 to 31, and differentiated progeny thereof.

33. An assay for the effect of presence in a cell of a protein or polypeptide or other product of DNA expression, comprising the steps:

(a) (i) transfecting the cell with a first vector that expresses a replication factor; or

(ii) otherwise obtaining a cell that expresses or will express the replication factor;

(b) transfecting the cell with a second vector, wherein

(i) the second vector contains a DNA coding for the protein or polypeptide or other product of DNA expression in operative combination with a promoter for expression of the DNA;

(ii) the second vector also contains a DNA coding for a selectable marker in operative combination with a promoter for expression of the selectable marker; and

(iii) extrachromosomal replication of the second vector is dependant

upon presence within the cell of the replication factor;

(c) selecting for cells that have been transfected with the second vector; and

(d) maintaining the selected cells over a plurality of generations so as to assay the effect of expression of the protein or polypeptide or other product of DNA expression.

34. An assay according to Claim 33 wherein step (a) is carried out once and the cells obtained are divided and used for a plurality of separate assays in which steps (b)-(d) are carried out a plurality of times with second vectors containing different DNA sequences.

35. An assay according to Claim 33 or 34 for assay of the effect of presence in the cell of two factors, each factor being independently selected from a protein, a polypeptide and another product of DNA expression.

36. A method of screening a library of cDNAs comprising assaying the effect of expression of each of the cDNAs according to the method of any of Claims 33 to 35.

37. A method of investigating the properties of a DNA sequence comprising expressing in a cell a composite DNA including (a) the DNA sequence under investigation, linked to (b) a DNA coding for a cell active protein, wherein

activity of the cell active protein is dependant upon transport of the cell active protein to the cell surface, and

the DNA of (b) does not code for a polypeptide capable of directing transportation of the cell active protein to the cell surface.

38. A method according to Claim 37 for screening a library of DNAs to identify DNA sequences coding for signal polypeptide sequences that transport

proteins to the cell surface, and the method optionally comprises determining whether the cell active protein is transported to the cell surface and remains there or is secreted by the cell.

39. A method according to Claim 37 or 38 wherein the DNA of (b) is obtained by deleting or disabling, from a DNA encoding a cell surface or secreted protein, that portion of the DNA that codes for the polypeptide sequence responsible for transportation of the protein to the cell surface.

40. A method according to any of Claims 37 to 39 wherein the cell active protein induces a morphological or proliferative change in the cell.

41. A method according to any of Claims 37 to 40 wherein the cell active protein inhibits differentiation of the cell and in the absence of the cell active protein the cell will differentiate.

42. A method according to any of Claims 37 to 41 wherein the cell active protein is a cell surface receptor.

43. A method according to Claim 42 wherein the cell active protein is an IL-6 receptor and the DNA of (b) encodes a modified form of the receptor preprotein lacking a functional signal sequence.

44. A method according to any of Claims 37 to 43 comprising investigating the properties of a DNA in mammalian or avian cells.

45. A method according to any of Claims 37 to 44 comprising investigating the properties of a DNA in embryonic cells.

46. A method according to Claim 45 comprising investigating the properties of a DNA in ES, EC or EG cells or differentiated progeny of such cells.

47. A method according to any of Claims 37 to 46 comprising expressing the composite DNA by:

(a) (i) transfecting the cell with a first vector that expresses a replication factor; or

(ii) otherwise obtaining a cell that expresses or will express the replication factor;

(b) transfecting the cell with a second vector, wherein

(i) the second vector contains the composite DNA in operative combination with a promoter for expression of the composite DNA;

(ii) the second vector also contains a DNA coding for a selectable marker in operative combination with a promoter for expression of the selectable marker; and

(iii) extrachromosomal replication of the second vector is dependant upon presence within the cell of the replication factor;

(c) selecting for cells that have been transfected with the second vector; and

(d) maintaining the selected cells over a plurality of generations so as to assay the effect of expression of the composite DNA.

48. A method according to claim 47 wherein step (a) is carried out once and the cells obtained are divided and used for a plurality of separate methods in which steps (b)-(d) are carried out a plurality of times with second vectors containing different DNA sequences.

49. A method according to any of Claims 37 to 48 for identification of a DNA coding for a cell surface or secreted protein.

50. A method according to any of Claims 37 to 48 for identification of a cell surface or secreted protein.

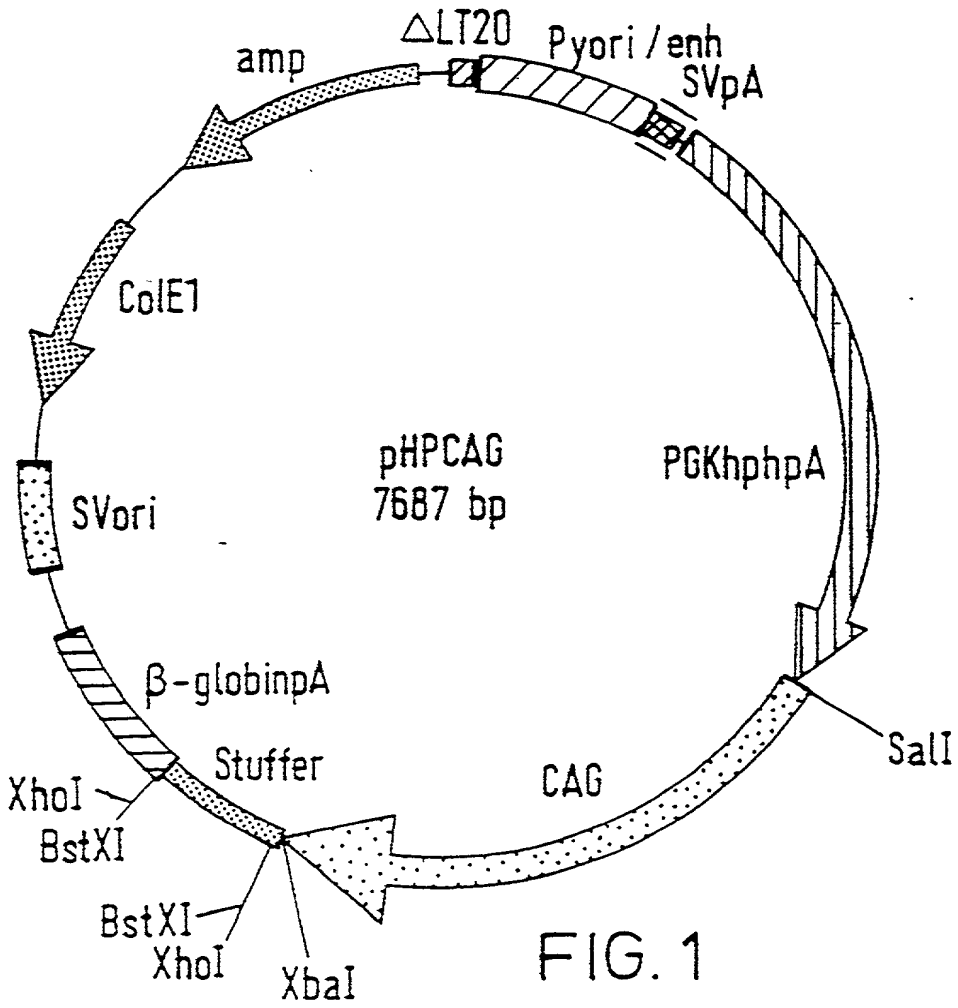
amp
ΔLT20
Pyori/enh
SVpA
ColE1
pHPCAG
7687 bp
PGKhphpA
SVori
β-globinpA
Stuffer
CAG
SalI
XhoI
BstXI
BstXI
XhoI
XbaI

FIG. 1

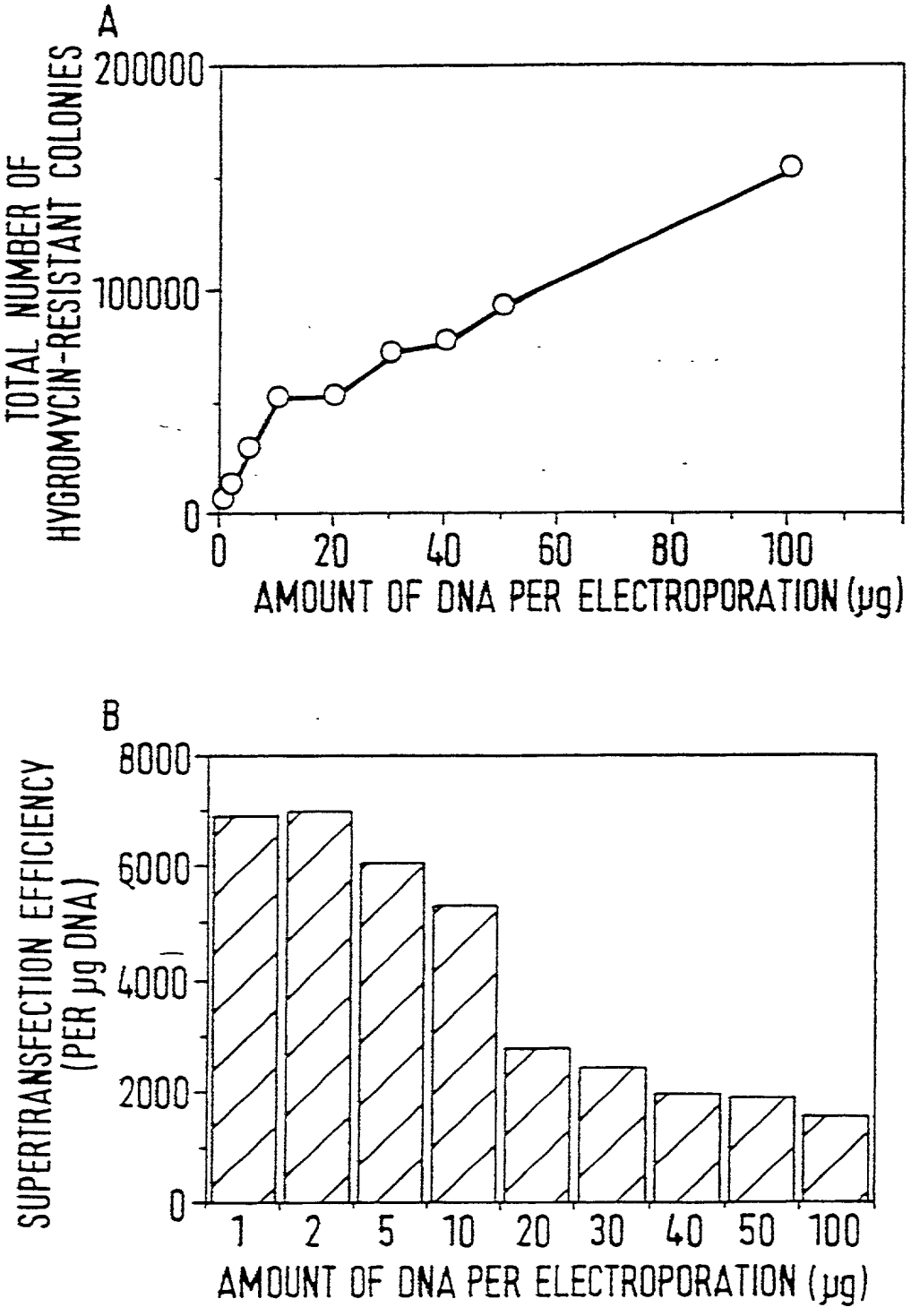
A
TOTAL NUMBER OF HYGROMYCIN-RESISTANT COLONIES
200000
100000
0
0
20
40
60
80
100
AMOUNT OF DNA PER ELECTROPORATION (µg)
B
SUPERTRANSFECTION EFFICIENCY (PER µg DNA)
8000
6000
4000
2000
0
1
2
5
10
20
30
40
50
100
AMOUNT OF DNA PER ELECTROPORATION (µg)

FIG. 2

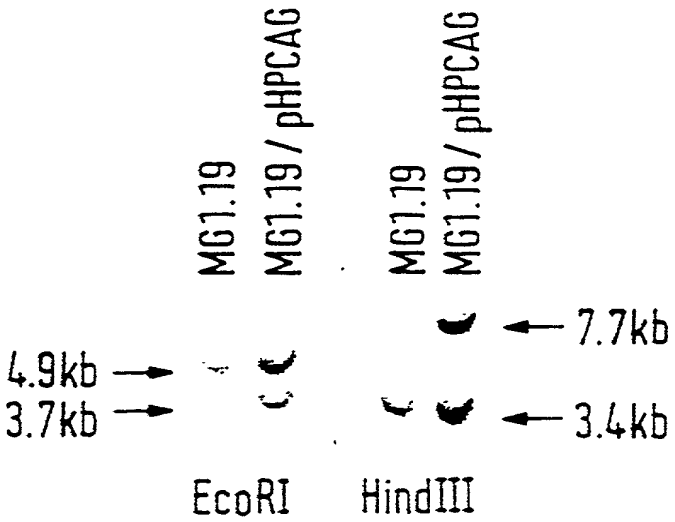
MG1.19
MG1.19 / pHPCAG
MG1.19
MG1.19 / pHPCAG
7.7kb
4.9kb
3.7kb
3.4kb
EcoRI
HindIII

FIG. 3

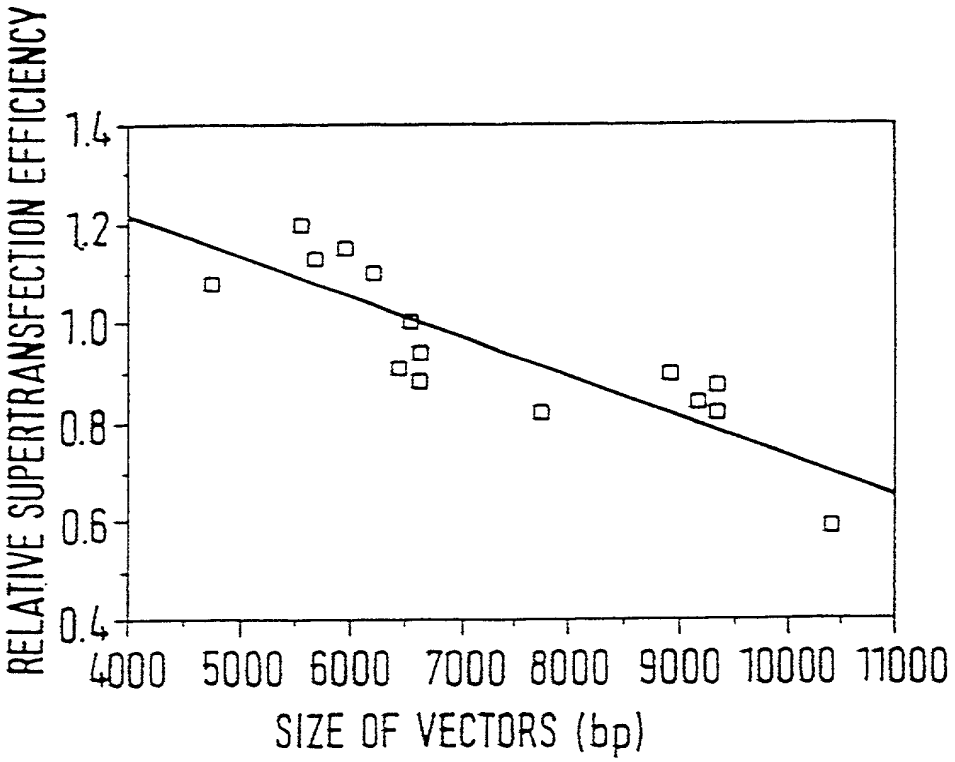
RELATIVE SUPERTRANSFECTION EFFICIENCY
1.4
1.2
1.0
0.8
0.6
0.4
4000
5000
6000
7000
8000
9000
10000
11000
SIZE OF VECTORS (bp)

FIG. 4

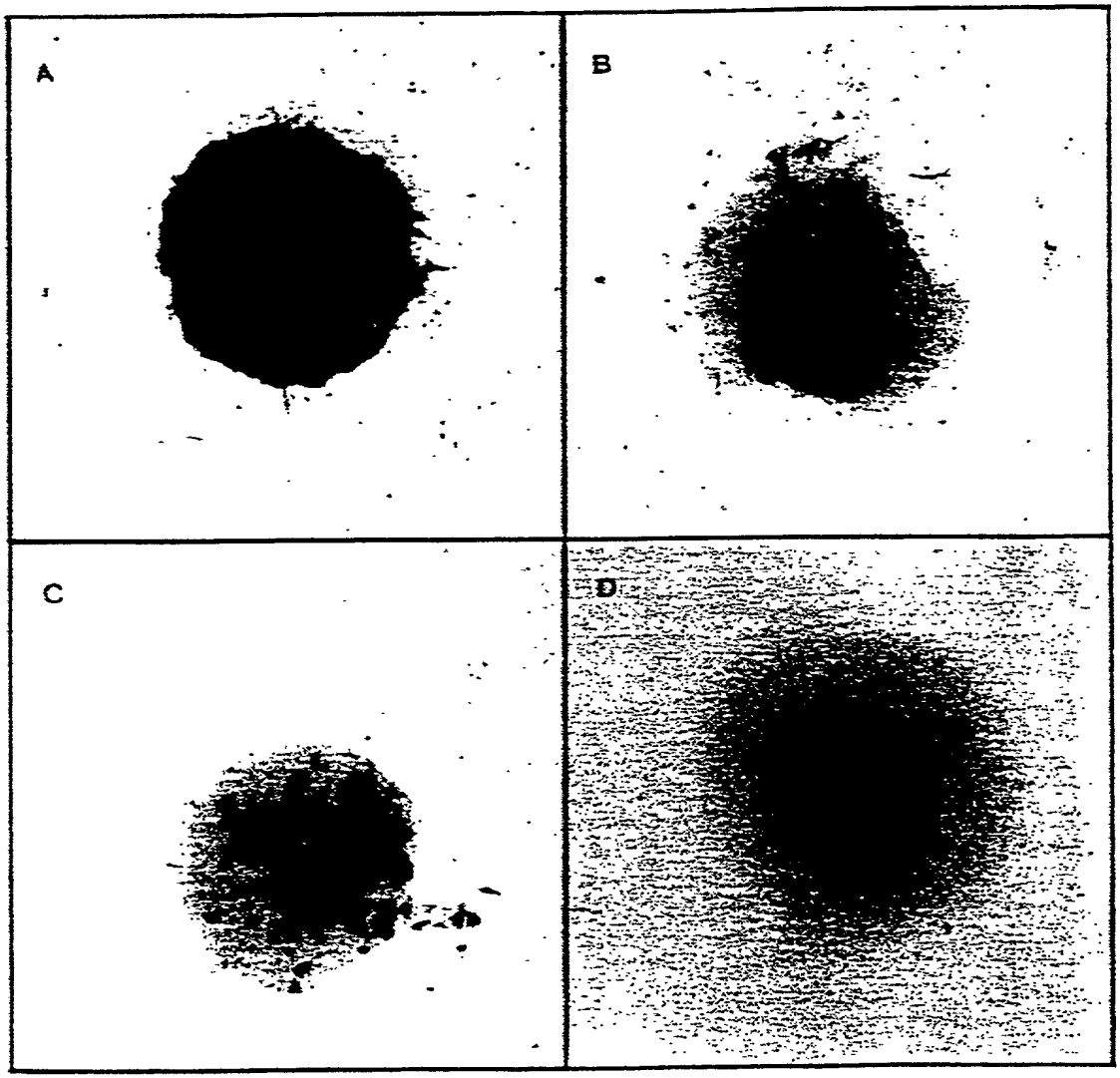

FIG. 5

FIG. 6

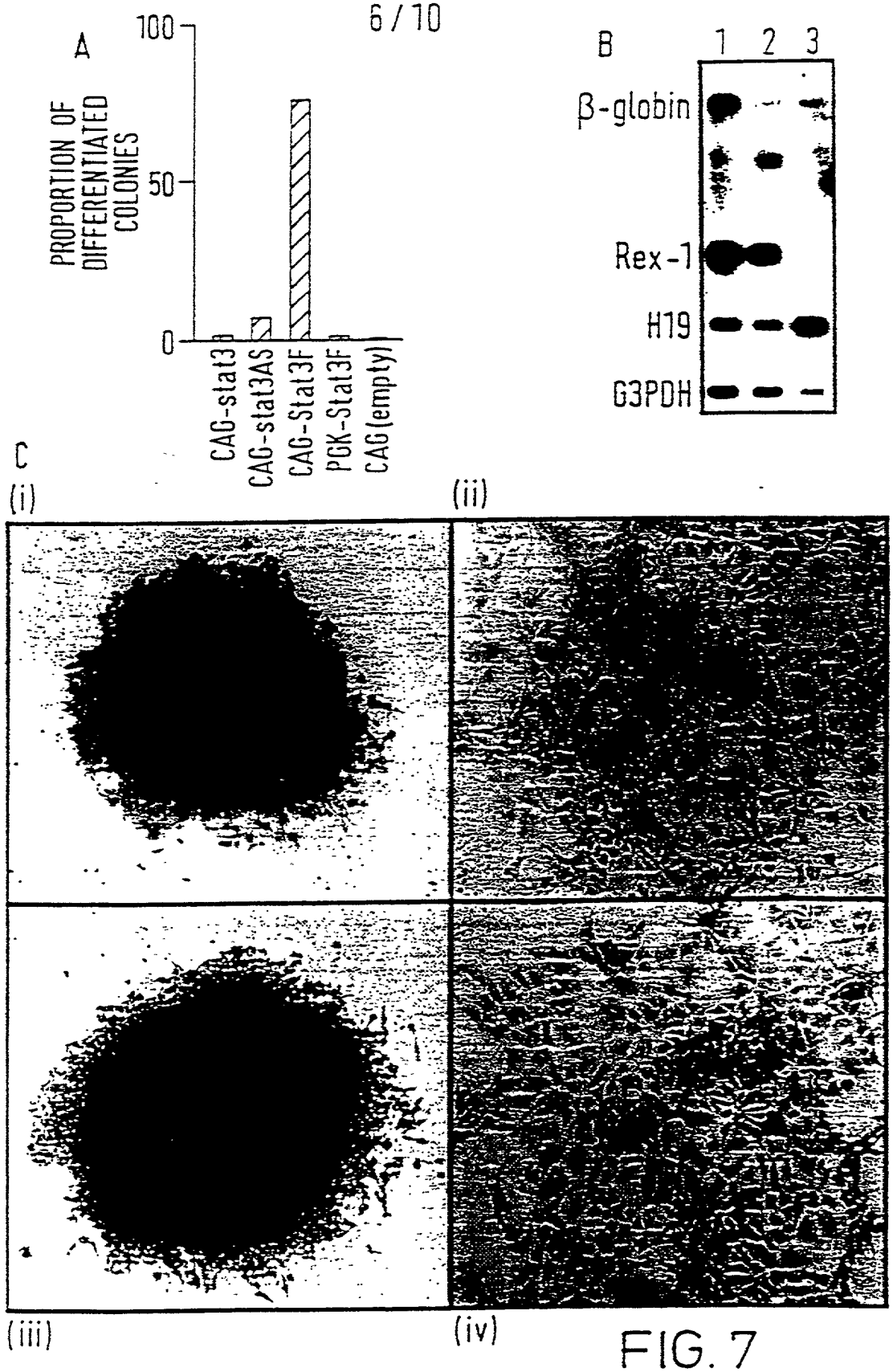
A
100
50
0
PROPORTION OF DIFFERENTIATED COLONIES
CAG-stat3
CAG-stat3AS
CAG-Stat3F
PGK-Stat3F
CAG(empty)
B
1 2 3
β-globin
Rex-1
H19
G3PDH
C
(i)
(ii)
(iii)
(iv)

FIG. 7

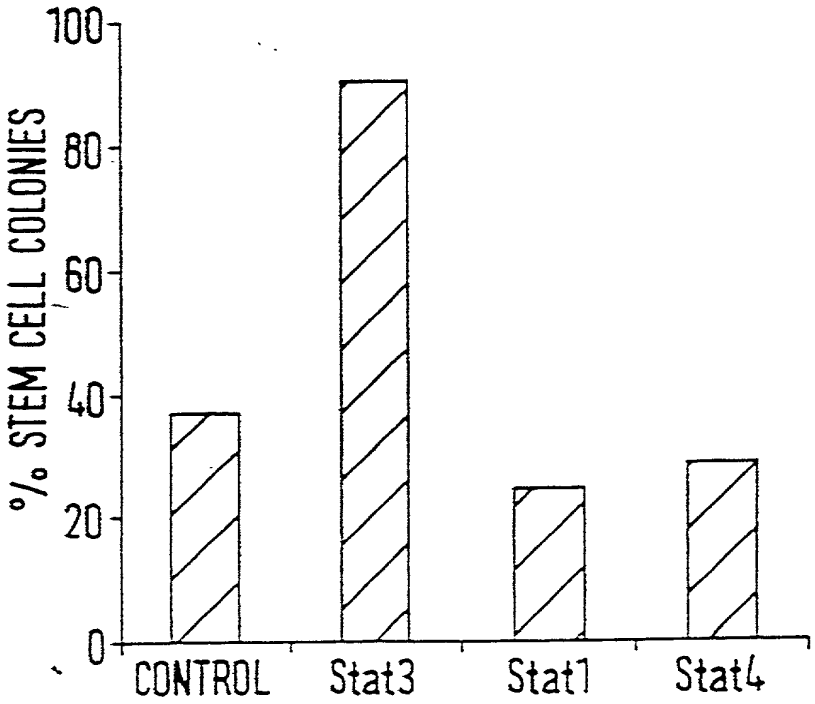
100
80
60
40
20
0
% STEM CELL COLONIES
CONTROL
Stat3
Stat1
Stat4

FIG. 8

FIG. 9

SEQUENCE OF LINKER OLIGONUCLEOTIDES.

a) FLAG linker

```
                                                    FLAG epitope
          XhoI          NotI     Gly Ser ASP TYR LYS ASP ASP
ASP  CTA GAC TCG AGT AGC GGC CGC GGC AGC GAC TAC AAG GAC
     GAC GAC

                          BssHII
ASP LYS Gly Ser Cys Arg Ala
GAC AAG GGG AGC TGC CGC GCG C
```

b) [gly_4 ser]$_2$ linker

```
          XhoI          NotI     Gly Gly Gly Gly Ser Gly Gly
Gly Gly Ser
CTA GAC TCG AGT AGC GGC CGC GGA GGC GGA TAC GGA AGC GGA
GGA GGG AGC

   BssHII
Cys Arg Ala
TGC CGC GCG C
```

FIG. 10

SEQUENCE OF TRUNCATED AND MODIFIED IL6R.

a) FLAGdeltaIL6R

TCTAGACTCGAGTAGCGGCCGCGGCAGCGACTACAAGGACGACGACGACAAGGGGAGCTGCCGCGCGCTGGAGGT
GGCAAATGGCACAGTGACAAGCCTGCCAGGGGCCACCGTTACCCTGATTTGCCCCGGGAAGGAAGCAGCAGGCAA
TGTTACCATTCACTGGGTGTACTCTGGCTCACAAAACAGAGAATGGACTACCACAGGAAACACACTGGTTCTGAG
GGACGTGCAGCTCAGCGACACTGGGGACTATTTATGCTCCCTGAATGATCACCTGGTGGGGACTGTGCCCTTGCT
GGTGGATGTTCCCCCAGAGGAGCCCAAGCTCTCCTGCTTCCGGAAGAACCCCCTTGTCAACGCCATCTGTGAGTG
GCGTCCGAGCAGCACCCCCTCTCCAACCACGAAGGCTGTGCTGTTTGCAAAGAAAATCAACACCACCAACGGGAA
GAGTGACTTCCAGGTGCCCTGCCAGTATTCTCAGCAGCTGAAAAGCTTCTCCTGCCAGGTGGAGATCCTGGAGGG
TGACAAAGTATACCACATAGTGTCACTGTGCGTTGCAAACAGTGTGGGAAGCAAGTCCAGCCACAACGAAGCGTT
TCACAGCTTAAAAATGGTGCAGCCGGATCCACCTGCCAACCTTGTGGTATCAGCCATACCTGGAAGGCCGCGCTG
GCTCAAAGTCAGCTGGCAGCACCCTGAGACCTGGGACCCGAGTTACTACTTGCTGCAGTTCCAGCTTCGATACCG
ACCTGTATGGTCAAAGGAGTTCACGGTGTTGCTGCTCCCGGTGGCCCAGTACCAATGCGTCATCCATGATGCCTT
GCGAGGAGTGAAGCACGTGGTCCAGGTCCGTGGGAAGGAGGAGCTTGACCTTGGCCAGTGGAGTGAATGGTCCCC
AGAGGTCACGGGCACTCCTTGGATAGCAGAGCCCAGGACCACCCCGGCAGGAATCCTCTGGAACCCCACACAGGT
CTCTGTTGAAGACTCTGCCAACCACGAGGATCAGTACGAAAGTTCTACAGAAGCAACGAGTGTCCTCGCCCCAGT
GCAAGAATCCTCGTCCATGTCCCTGCCCACATTCCTGGTAGCTGGAGGAAGCTTGGCGTTTGGGTTGCTTCTCTG
TGTCTTCATCATCCTGAGACTCAAGCAGAAATGGAAGTCAGAGGCTGAGAAGGAAAGCAAGACGACCTCTCCTCC
ACCCCCACCGTATTCCTTGGGCCCACTGAAGCCGACCTTCCTTCTGGTTCCTCTCCTCACCCCACACAGCTCTGG
GTCTGACAATACCGTAAACCACAGCTGCCTGGGTGTCAGGGACGCACAGAGCCCTTATGACAACAGCAACAGAGA
CTACTTATTCCCCAGATAA

b)[gly$_4$ser]$_2$deltaIL6R

TCTAGACTCGAGTAGCGGCCGCGGAGGCGGAGGAAGCGGAGGAGGAGGGAGCTGCCGCGCGCTGGAGGTGGCAAA
TGGCACAGTGACAAGCCTGCCAGGGGCCACCGTTACCCTGATTTGCCCCGGGAAGGAAGCAGCAGGCAATGTTAC
CATTCACTGGGTGTACTCTGGCTCACAAAACAGAGAATGGACTACCACAGGAAACACACTGGTTCTGAGGGACGT
GCAGCTCAGCGACACTGGGGACTATTTATGCTCCCTGAATGATCACCTGGTGGGGACTGTGCCCTTGCTGGTGGA
TGTTCCCCCAGAGGAGCCCAAGCTCTCCTGCTTCCGGAAGAACCCCCTTGTCAACGCCATCTGTGAGTGGCGTCC
GAGCAGCACCCCCTCTCCAACCACGAAGGCTGTGCTGTTTGCAAAGAAAATCAACACCACCAACGGGAAGAGTGA
CTTCCAGGTGCCCTGCCAGTATTCTCAGCAGCTGAAAAGCTTCTCCTGCCAGGTGGAGATCCTGGAGGGTGACAA
AGTATACCACATAGTGTCACTGTGCGTTGCAAACAGTGTGGGAAGCAAGTCCAGCCACAACGAAGCGTTTCACAG
CTTAAAAATGGTGCAGCCGGATCCACCTGCCAACCTTGTGGTATCAGCCATACCTGGAAGGCCGCGCTGGCTCAA
AGTCAGCTGGCAGCACCCTGAGACCTGGGACCCGAGTTACTACTTGCTGCAGTTCCAGCTTCGATACCGACCTGT
ATGGTCAAAGGAGTTCACGGTGTTGCTGCTCCCGGTGGCCCAGTACCAATGCGTCATCCATGATGCCTTGCGAGG
AGTGAAGCACGTGGTCCAGGTCCGTGGGAAGGAGGAGCTTGACCTTGGCCAGTGGAGTGAATGGTCCCCAGAGGT
CACGGGCACTCCTTGGATAGCAGAGCCCAGGACCACCCCGGCAGGAATCCTCTGGAACCCCACACAGGTCTCTGT
TGAAGACTCTGCCAACCACGAGGATCAGTACGAAAGTTCTACAGAAGCAACGAGTGTCCTCGCCCCAGTGCAAGA
ATCCTCGTCCATGTCCCTGCCCACATTCCTGGTAGCTGGAGGAAGCTTGGCGTTTGGGTTGCTTCTCTGTGTCTT
CATCATCCTGAGACTCAAGCAGAAATGGAAGTCAGAGGCTGAGAAGGAAAGCAAGACGACCTCTCCTCCACCCCC
ACCGTATTCCTTGGGCCCACTGAAGCCGACCTTCCTTCTGGTTCCTCTCCTCACCCCACACAGCTCTGGGTCTGA
CAATACCGTAAACCACAGCTGCCTGGGTGTCAGGGACGCACAGAGCCCTTATGACAACAGCAACAGAGACTACTT
ATTCCCCAGATAA

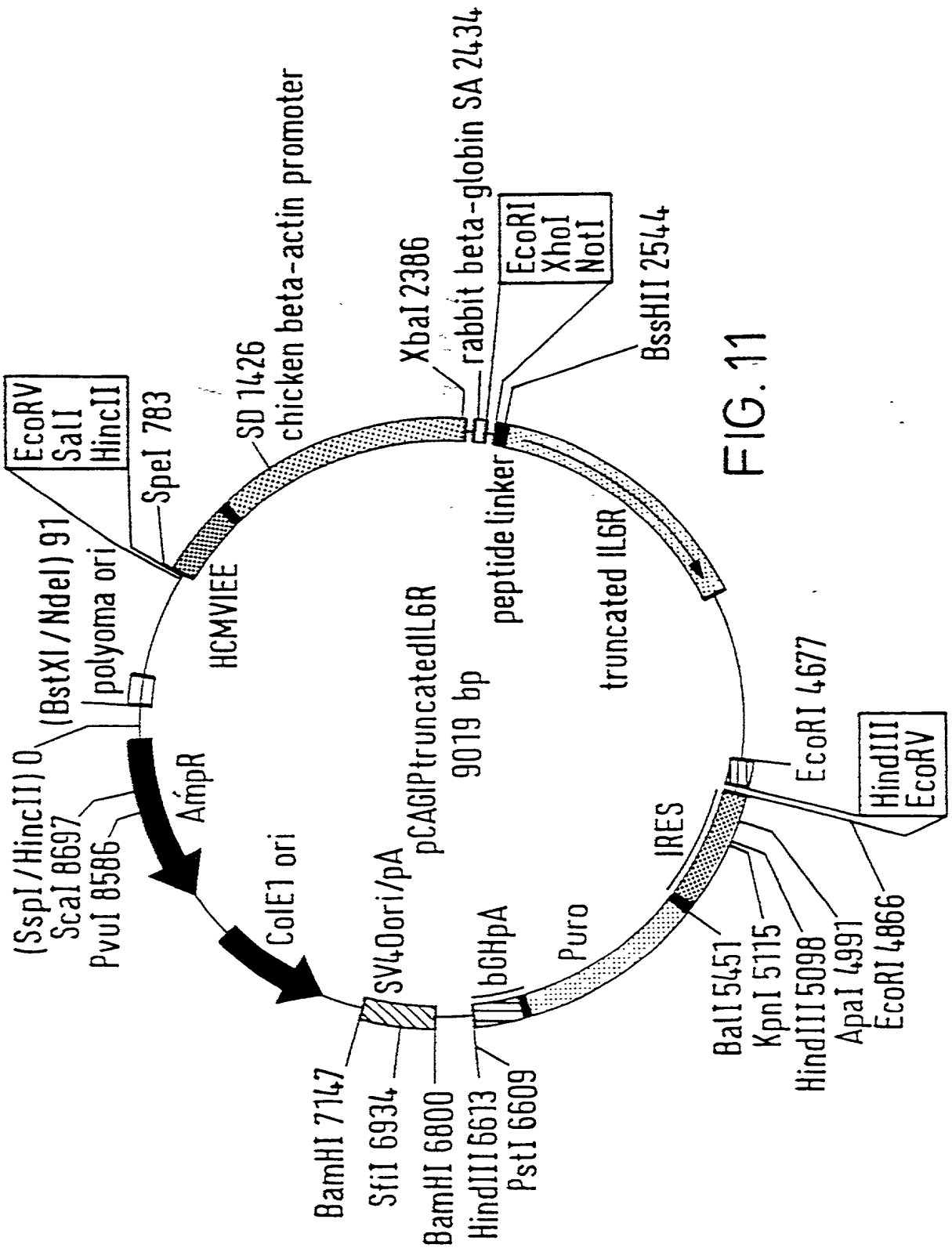
EcoRV
SalI
HincII
SpeI 783
SD 1426
chicken beta-actin promoter
XbaI 2386
rabbit beta-globin SA 2434
EcoRI
XhoI
NotI
BssHII 2544
(BstXI / NdeI) 91
polyoma ori
HCMVIEE
peptide linker
truncated IL6R
(SspI / HincII) 0
ScaI 8697
PvuI 8586
AmpR
pCAGIPtruncatedIL6R
9019 bp
EcoRI 4677
HindIII
EcoRV
ColE1 ori
SV40ori/pA
bGHpA
Puro
IRES
BamHI 7147
SfiI 6934
BamHI 6800
HindIII 6613
PstI 6609
BalI 5451
KpnI 5115
HindIII 5098
ApaI 4991
EcoRI 4866

FIG. 11

DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that: my residence, post office address and citizenship are as stated below next to my name; I believe I am the original, first, and sole inventor (if only one name is listed below) or an original, first, and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

DNA EXPRESSION IN TRANSFECTED CELLS AND ASSAYS CARRIED OUT IN TRANSFECTED CELLS the specification of which ☒ is attached and/or ☐ was filed on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above. I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR § 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. § 119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate or § 365(a) of any PCT international application(s) designating at least one country other than the United States, listed below and have also identified below, any foreign application(s) for patent or inventor's certificate, or any PCT International application(s) having a filing date before that of the application(s) of which priority is claimed:

Country	Application Number	Date of Filing	Priority Claimed Under 35 U.S.C.
GREAT BRITAIN	9701492.2	24 January 1997	☒ YES ☐ NO
			☐ YES ☐ NO

I hereby claim the benefit under 35 U.S.C. § 119(e) of any United States provisional application(s) listed below:

Application Number	Date of Filing

I hereby claim the benefit under 35 U.S.C. § 120 of any United States application(s) or § 365(c) of any PCT International application(s) designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application(s) in the manner provided by the first paragraph of 35 U.S.C. § 112, I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR § 1.56 which became available between the filing date of the prior application(s) and the national or PCT International filing date of this application:

Application Number	Date of Filing	Status (Patented, Pending, Abandoned)
PCT/GB98/00216	26 January 1998	Pending

I hereby appoint the following attorney and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith. FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, L.L.P., Douglas B. Henderson, Reg. No. 20,291; Ford F. Farabow, Jr., Reg. No. 20,630; Arthur S. Garrett, Reg. No. 20,338; Donald R. Dunner, Reg. No. 19,073; Brian G. Brunsvold, Reg. No. 22,593; Tipton D. Jennings, IV, Reg. No. 20,645; Jerry D. Voight, Reg. No. 23,020; Laurence R. Hefter, Reg. No. 20,827; Kenneth E. Payne, Reg. No. 23,098; Herbert H. Mintz, Reg. No. 26,691; C. Larry O'Rourke, Reg. No. 26,014; Albert J. Santorelli, Reg. No. 22,610; Michael C. Elmer, Reg. No. 25,857; Richard H. Smith, Reg. No. 20,609; Stephen L. Peterson, Reg. No. 26,325; John M. Romary, Reg. No. 26,331; Bruce C. Zotter, Reg. No. 27,680; Dennis P. O'Reilley, Reg. No. 27,932; Allen M. Sokal, Reg. No. 26,695; Robert D. Bajefsky, Reg. No. 25,387; Richard L. Stroup, Reg. No. 28,478; David W. Hill, Reg. No. 28,220; Thomas L. Irving, Reg. No. 28,619; Charles E. Lipsey, Reg. No. 28,165; Thomas W. Winland, Reg. No. 27,605; Basil J. Lewris, Reg. No. 28,818; Martin I. Fuchs, Reg. No. 28,508; E. Robert Yoches, Reg. No. 30,120; Barry W. Graham, Reg. No. 29,924; Susan Haberman Griffen, Reg. No. 30,907; Richard B. Racine, Reg. No. 30,415; Thomas H. Jenkins, Reg. No. 30,857; Robert E. Converse, Jr., Reg. No. 27,432; Clair X. Mullen, Jr., Reg. No. 20,348; Christopher P. Foley, Reg. No. 31,354; John C. Paul, Reg. No. 30,413; Roger D. Taylor, Reg. No. 28,992; David M. Kelly, Reg. No. 30,953; Kenneth J. Meyers, Reg. No. 25,146; Carol P. Einaudi, Reg. No. 32,220; Walter Y. Boyd, Jr., Reg. No. 31,738; Steven M. Anzalone, Reg. No. 32,095; Jean B. Fordis, Reg. No. 32,984; Barbara C. McCurdy, Reg. No. 32,120; James K. Hammond, Reg. No. 31,964; Richard V. Burgujian, Reg. No. 31,744; J. Michael Jakes, Reg. No. 32,824; Dirk D. Thomas, Reg. No. 32,600; Thomas W. Banks, Reg. No. 32,719; Christopher P. Isaac, Reg. No. 32,616; Bryan C. Diner, Reg. No. 32,409; M. Paul Barker, Reg. No. 32,013; Andrew Chanho Sonu, Reg. No. 33,457; David S. Forman, Reg. No. 33,694; Vincent P. Kovalick, Reg. No. 32,867; James W. Edmondson, Reg. No. 33,871; Michael R. McGurk, Reg. No. 32,045; Joann M. Neth, Reg. No. 36,363; Gerson S. Panitch, Reg. No. 33,751; Cheri M. Taylor, Reg. No. 33,216; Charles E. Van Horn, Reg. No. 40,266; Linda A. Wadler, Reg. No. 33,218; Jeffrey A. Berkowitz, Reg. No. 36,743; Michael R. Kelly, Reg. No. 33,921; and James B. Monroe, Reg. No. 33,971; and Leslie A. McDonell, Reg. No. 34,872. Please address all correspondence to FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, L.L.P. 1300 I Street, N.W., Washington, D.C. 20005, Telephone No. (202) 408-4000.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Listing of Inventors Continued on Page 2 hereof. ☒ Yes ☐ No

Full Name of First Inventor Catherine Clare BLACKBURN	**Inventor's Signature**	**Date**
Residence Edinburgh GREAT BRITAIN		**Citizenship** Great Britain
Post Office Address The University of Edinburgh, Centre for Genome Research, The Kings Buildings, West Mains Road, Edinburgh EH9 3JQ Great Britain		

Full Name of Second Inventor Ian Paul CHAMBERS	**Inventor's Signature**	**Date**
Residence Edinburgh GREAT BRITAIN		**Citizenship** Great Britain
Post Office Address The University of Edinburgh, Centre for Genome Research, The Kings Buildings, West Mains Road, Edinburgh EH9 3JQ Great Britain		

Full Name of Third Inventor Alexander L. MEDVINSKY	**Inventor's Signature**	**Date**
Residence Edinburgh GREAT BRITAIN		**Citizenship** Russia
Post Office Address The University of Edinburgh, Centre for Genome Research, The Kings Buildings, West Mains Road, Edinburgh EH9 3JQ Great Britain		

Full Name of Fourth Inventor Hitoshi NIWA	**Inventor's Signature**	**Date**
Residence Edinburgh GREAT BRITAIN		**Citizenship** Japan
Post Office Address The University of Edinburgh, Centre for Genome Research, The Kings Buildings, West Mains Road, Edinburgh EH9 3JQ Great Britain		

Full Name of Fifth Inventor Austin G. SMITH	**Inventor's Signature**	**Date**
Residence Edinburgh GREAT BRITAIN		**Citizenship** Great Britain
Post Office Address The University of Edinburgh, Centre for Genome Research, The Kings Buildings, West Mains Road, Edinburgh EH9 3JQ Great Britain		

Full Name of Sixth Inventor	**Inventor's Signature**	**Date**
Residence		**Citizenship**
Post Office Address		

Full Name of Seventh Inventor	**Inventor's Signature**	**Date**
Residence		**Citizenship**
Post Office Address		